The

PHILOSOPHICAL MAGAZINE

FIRST PUBLISHED IN 1798

.. 44 SEVENTH SERIES No. 356 September 1953

A Journal of Theoretical Experimental and Applied Physics

PRICE 15s. 0d.
Annual Subscription £8 0s. 0d. payable in advance

D AND PUBLISHED BY TAYLOR & FRANCIS LTD., RED LION COURT, FLEET ST., LONDON, E.C.4.

Early Scientific Publications

DIARY OF ROBERT HOOKE, M.A., M.D., F.R.S.
1672-1680

Edited by **H. W. ROBINSON** and **W. ADAMS**
Recommended for publication by the Royal Society, London

25/-
net

"This vivid record of the scientific, artistic and social activities of a remarkable man during remarkable years has too long remained in obscurity."—Extract from foreword by Sir Frederick Gowland Hopkins, O.M., President of the Royal Society.

MATHEMATICAL WORK OF JOHN WALLIS, D.D., F.R.S.

By **J. F. SCOTT, Ph.D., B.A.**

12/6
net

"His work will be indispensable to those interested in the early history of The Royal Society. I commend to all students of the Seventeenth Century, whether scientific or humane, this learned and lucid book."—Extract from foreword by Prof. E. N. da C. Andrade, D.Sc., Ph.D., F.R.S.
Recommended for publication by University of London

CORRESPONDENCE AND PAPERS OF EDMOND HALLEY

Arranged and Edited by **EUGENE FAIRFIELD MACPIKE**

21/-
net

First published on behalf of The History of Science Society by Oxford University Press. Now re-issued by Taylor & Francis, Ltd.

MEMOIRS OF SIR ISAAC NEWTON'S LIFE

5/-
net

By **WILLIAM STUKELEY, M.D., F.R.S., 1752**
From an Original Manuscript
Now in the possession of the Royal Society, London

HEVELIUS, FLAMSTEED AND HALLEY

12/6
net

Three Contemporary Astronomers and their Mutual Relations
By **EUGENE FAIRFIELD MACPIKE**
Published by arrangement with The History of Science Society

Established over 150 years

TAYLOR & FRANCIS, LTD.
RED LION COURT, FLEET ST., LONDON E.C.
PRINTERS & PUBLISHERS OF SCIENTIFIC BOOKS

XCVIII. *The Growth of Fatigue Cracks*

By A. K. Head
H. H. Wills Physical Laboratory, University of Bristol, and Aeronautical Research Laboratories, Melbourne*

[Received May 20, 1953]

Summary

The qualitative features of the growth of a fatigue crack are discussed. A model is proposed which has these features and from which quantitative predictions can be made. These are in satisfactory agreement with experimental measurements. It is concluded that cracks may be formed at an early stage in fatigue and that the growth of a crack may occupy a large part of the measured fatigue life.

§ 1. Introduction

In a review of the mechanism of fatigue of metals (Head 1953) it was suggested that there are three stages involved. The first stage is deformation throughout the bulk of the metal until it has work-hardened to the level of the applied stress. For fatigue stresses below the static yield point of the metal this first stage is absent. During the second stage isolated regions show slowly increasing deformation. It is in these regions that a crack is ultimately observed, the growth of which forms the final stage of failure by fatigue. A crack may become visible after 90–95% of the total number of cycles of stress required to produce complete fracture. This, together with the feeling that the high stress concentration at the tip of a crack will cause it to extend rapidly has led to the view that the middle stage forms the main part of the process of fatigue. However, as has been pointed out in the above review, there is a small amount of evidence that a crack can be formed early and grow very slowly, being present in some cases during 90% of the fatigue life of a specimen.

If this is the case, the mechanism of growth of a fatigue crack is of as great interest as the processes leading to the formation of the crack. In § 2 the qualitative features of the growth of a crack are considered and in the following sections a model of a crack is proposed which possesses these features and from which quantitative predictions can be made. In § 7 these predictions are compared with experiment and in § 8 some consequences are discussed.

§ 2.

It has been shown by a number of workers (Head 1953) that the average direction of extension of a fatigue crack is generally perpendicular to the direction of maximum tension-compression. There are local

* Communicated by the Author.

deviations which are governed by the crystalline structure of the metal, but this factor will not be considered here.

Consider then a crack in the form of a thin elliptic cylindrical hole in a uniform material and the consequences of applying a tensile stress perpendicular to the length of the crack (fig. 1).

If the material is purely elastic, then there is a large stress concentration at the tip of the crack (Inglis 1913) and the stress acting across the plane $y=0$ ahead of the crack will be of the form shown diagrammatically in fig. 2 by the dashed curve.

Fig. 1

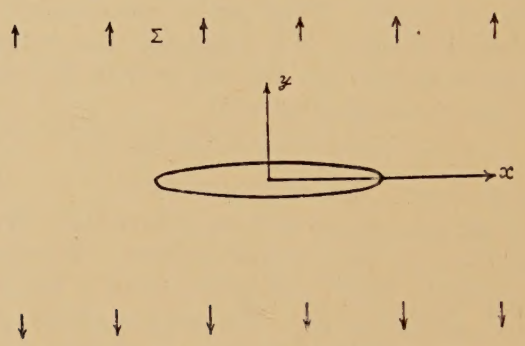
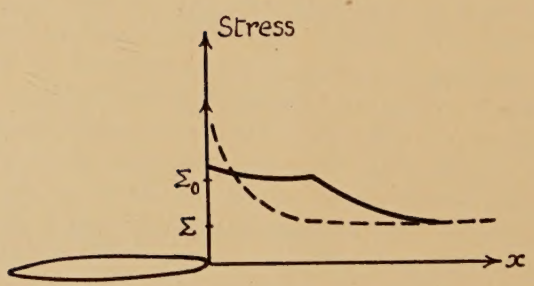


Fig. 2



For a material which yields plastically above a certain stress, Σ_0 say, a plastic region will be formed round the tip of the crack. In this case the stress distribution will be governed by three effects :

(i) As the plastic region yields, it work-hardens.

(ii) The plastic region is constrained by the elastic surroundings, which, as yielding occurs, take up some of the load, thus reducing the stress applied to the plastic region.

(iii) This increase in stress in the surrounding regions will cause some parts, which in the purely elastic case carried a stress less than Σ_0, to be subjected to a large enough stress to become plastic. That is, the plastic region extends further out than the point where the elastic stress would be equal to Σ_0.

In this case the stress distribution ahead of the tip of the crack would be of the form shown by the solid curve in fig. 2.

If the tensile stress is removed and an equal compressive stress applied, then plastic flow in the opposite sense will occur, again governed by the above three effects. If repeated cycles of tensile and compressive stresses are applied, then one of two situations will be reached ultimately (Orowan 1939). If the applied stress or the stress concentration is small enough, then the material round the tip of the crack may work-harden sufficiently to support the stress there by purely elastic deformations. But if this stress at the tip of the crack under purely elastic deformation is too large, it may be beyond the capacity of the material to work-harden to this extent without fracturing. In this case after a number of stress cycles a small region at the tip of the crack will fracture, the length of the crack will be slightly increased, the stress concentration will increase so the region at the new tip of the crack will work-harden to the fracture stress in a smaller number of cycles and so the crack extends at an increasing rate.

§ 3.

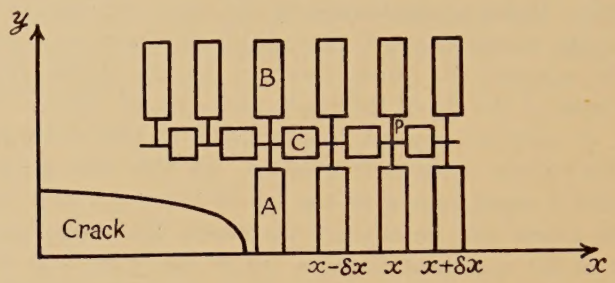
The process of § 2 by which a fatigue crack is visualized to extend is too complicated for solution by the present methods of mathematical plasticity. So a model of the process is proposed which is simplified sufficiently to be tractable yet nevertheless retains the essential features stated in § 2. The model is constructed in the following manner : the problem is first reduced to two dimensions by considering a crack in the form of a thin cylindrical cavity of infinite extent in the z-direction. The x and y axes are as indicated in fig. 1. Where necessary below, quantities refer to unit depth of the crack in the z-direction.

A one-dimensional model of this crack is constructed thus : a thin slice of the medium bounded by $y=\pm a_0$ is considered, which contains the crack and the plastic region at the tip of the crack, and through which the crack will extend. The medium is then divided into infinitesimal slices perpendicular to the x-axis and idealized into three types of mechanical elements. By symmetry only one quadrant of the crack need be considered and this dissection of the medium is shown in fig. 3 where neighbouring elements have been separated for clarity.

The elements A represent the region which becomes plastic, and fractures after a number of stress cycles. These are regarded as small tensile elements, of initial length a_0 with the initial stress–strain curve shown in fig. 4. Thus they are initially rigid up to stress Σ_0 after which they work-harden linearly until they fracture at stress Σ_1. The work-hardening modulus is denoted by F so the fracture strain is $(\Sigma_1-\Sigma_0)/F$. Due to symmetry their lower ends can be considered fixed along the line $y=0$. It is also assumed that these elements do not show any Bauschinger effect under reversed stressing. The consequences of this last point are discussed more fully below.

Elements B represent the constraining effect of the surrounding elastic material. These are taken as elastic tensile elements, of length d and Young's modulus E. Thus if any particular element A yields a distance ϵ the stress applied to it through the corresponding B element is reduced to $\Sigma - \epsilon E/d$.

Fig. 3



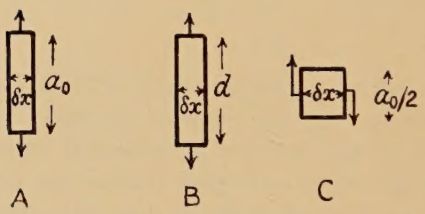
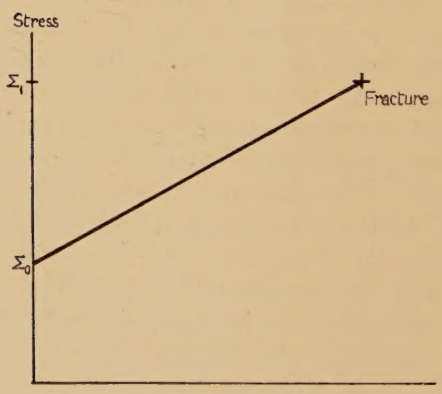


Fig. 4



Elements C provide the effect noted in (iii) of § 2, that if a particular A element yields it throws more stress on the neighbouring A elements. These are regarded as shear elements, of shear modulus G (which as a sufficient approximation we take equal to $\frac{1}{2}E$) and dimensions shown in fig. 3. Their length is taken as $\frac{1}{2}a_0$ because this coupling between adjacent A elements is really spread out over the interface between them and this factor compensates for concentrating the coupling into one element at the top end of A.

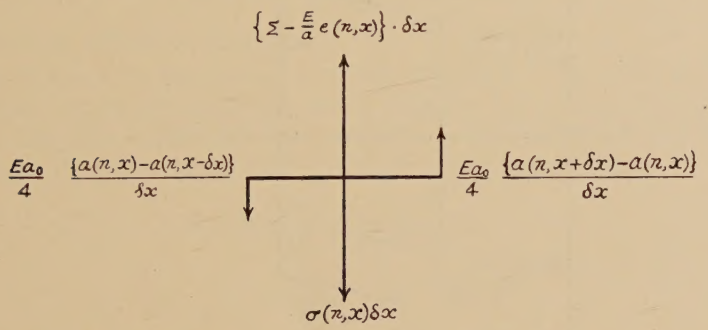
We now consider the response of this model to alternate applications of stresses $\pm\Sigma$ (which is assumed to be less than the virgin yield point of the material). As the application of a tensile followed by a compressive stress is usually taken to be one stress cycle we will denote the tensile half cycles by n, $n+1$, ... and the compressive half cycles by $n+\frac{1}{2}$, $n+\frac{3}{2}$, ... On any particular application of load there are three separate regions to be considered :

(i) In the region of the crack there are only B and C elements active. It is convenient however to consider as being present A elements of zero strength.

(ii) Ahead of the crack there is a region in which the A elements yield during this half cycle.

(iii) Beyond this is an infinite region in which the stress is not large enough to cause yielding.

Fig. 5



We denote by $\sigma(n, x)$ the stress in the A element at a point x on the nth (tensile) loading. Then $\sigma(n, x)=0$ in region (i) and is less than Σ_0 in region (iii). On the $(n+\frac{1}{2})$th compressive loading we define $\sigma(n+\frac{1}{2}, x)$ to be the magnitude of the stress in the A element at x.

Denote by $a(n, x)$ the length of an A element on either a tensile or compressive loading. Then $a(n, x)=a_0$ in region (iii). It is convenient to define

$$\epsilon(n, x)=a(n, x)-a_0,$$

$$\epsilon(n+\tfrac{1}{2}, x)=a_0-a(n+\tfrac{1}{2}, x).$$

Figure 5 shows the forces acting at P on the nth loading. Equating these forces and letting $\delta x \to 0$ gives

$$\frac{Ea_0}{4}\frac{d^2\epsilon(n, x)}{dx^2}+\Sigma-\frac{E}{d}\epsilon(n, x)-\sigma(n, x)=0, \quad . \ . \ . \quad (1)$$

and an identical equation is obtained for $\epsilon(n+\frac{1}{2}, x)$ under compressive loading.

In region (iii) (1) reduces to $\epsilon(n, x)=0$ and $\sigma(n, x)=\Sigma$. In region (i) $\sigma(n, x)=0$ so (1) becomes

$$\frac{Ea_0}{4}\frac{d^2\epsilon(n, x)}{dx^2}-\frac{E}{d}\epsilon(n, x)=-\Sigma. \quad . \ . \ . \ . \ . \quad (2)$$

Next consider the hysteresis loop of an A element in region (ii) (fig. 6), for which we have assumed the absence of any Bauschinger effect. Then

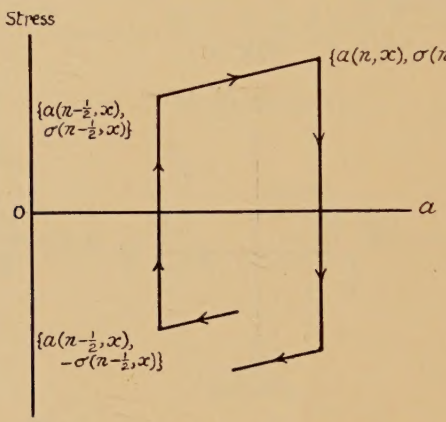
$$\sigma(n, x)-\sigma(n-\tfrac{1}{2}, x) = \frac{F}{a_0}\{a(n, x)-a(n-\tfrac{1}{2}, x)\}$$
$$= \frac{F}{a_0}\{\epsilon(n, x)+\epsilon(n-\tfrac{1}{2}, x)\}. \quad . \; . \quad (3)$$

Substituting $\sigma(n, x)$ from (3) in (1) gives, in region (ii)

$$\frac{Ea_0}{4}\frac{d^2\epsilon(n, x)}{dx^2} - \left(\frac{E}{d}+\frac{F}{a_0}\right)\epsilon(n, x) = -\Sigma+\sigma(n-\tfrac{1}{2}, x) + \frac{F}{a_0}\epsilon(n-\tfrac{1}{2}, x). \quad (4)$$

This equation is true for n integral or half integral.

Fig. 6



If the work-hardening per half cycle is assumed small, (2) and (4) can be written in the form

$$\frac{d^2\epsilon(n, x)}{dx^2} - k^2\epsilon(n, x) = -\frac{4}{Ea_0}\Sigma \qquad x<b(n) \quad . \; . \; . \; . \; . \quad (5)$$
$$= f(n-\tfrac{1}{2}, x) \qquad b(n)<x<c(n) \quad . \; . \; . \quad (6)$$

where
$$k^2 = \frac{4}{a_0 d} \quad . \; . \; . \; . \; . \; . \; . \; . \; . \; . \quad (7)$$

$$f(n-\tfrac{1}{2}, x) = \frac{4}{Ea_0}\left\{-\Sigma+\sigma(n-\tfrac{1}{2}, x)+\frac{F}{a_0}\epsilon(n-\tfrac{1}{2}, x)\right\}, \quad . \; . \quad (8)$$

and $b(n)$, $c(n)$ are the unknown positions of the boundaries between regions (i)–(ii) and (ii)–(iii) respectively, i.e. $b(n)$ is the current half length of the crack.

In the solution of (5), (6) the following boundary conditions must be satisfied :

at $x=0$: $\dfrac{d\epsilon(n, x)}{dx} = 0$ by symmetry about $x=0$, . . (9)

at $x=b(n)$: $\epsilon(n, x)$ and $[d\epsilon(n, x)]/dx$ must be continuous in order that the stresses at the boundary should remain finite. . . (10)

$\sigma(n, b(n))=\Sigma_1$ by definition of $b(n)$, ... (11)

at $x=c(n)$: $\epsilon(n, x)=0=[d\epsilon(n, x)]/dx$ for continuity, ... (12)

$\sigma(n, c(n))=\Sigma_0$ by definition of $c(n)$, ... (13)

at $n=0$: the initial state of the material to be specified. The rate of growth of the crack will be superposition of a steady state rate and a transient term depending upon the initial conditions. In solving for the steady state rate of growth the appropriate initial condition is that the material into which the crack will extend is uniform and with initial yield stress Σ_0.

§ 4. Solution in Region (i)

The solution of (5) satisfying (9) is

$$\epsilon(n, x) = \frac{d}{E}\Sigma + M \cosh kx, \quad . . . . . \quad (14)$$

where M is an arbitrary constant.

The constant d in our model is so far arbitrary. To make it definite we compare our model with the exact solution for a crack in an elastic medium. In our model this means that $\epsilon(n, x)=0$ at $x=b$ and by (14) the increase in width of the crack at $x=0$ under stress Σ is

$$\epsilon(n, 0) = \frac{d}{E}\Sigma(1-\operatorname{sech} kb)$$

$$= \frac{d}{E}\Sigma \quad \text{for} \quad kb \gg 1.$$

From the exact elastic solution for this case (Yoffe 1951, eqn. (8)) the corresponding increase in width is

$$\frac{b}{E}\Sigma . 2(1-\nu^2) \qquad \nu=\text{Poisson's ratio},$$

and so as a sufficient approximation take $d=b$, $2b$ being the total length of the crack.

§ 5. Solution in Region (ii)

The solution of (6) satisfying (12) is

$$\epsilon(n, x) = \frac{1}{k}\int_{c(n)}^{x} f(n-\tfrac{1}{2}, t) \sinh k(x-t)\, dt. \quad . . . \quad (15)$$

From the definition of f by (8)

$$f(n, x)-f(n-\tfrac{1}{2}, x) = \frac{4}{Ea_0}\left\{\sigma(n, x)-\sigma(n-\tfrac{1}{2}, x) + \frac{F}{a_0}(\epsilon(n, x)-\epsilon(n-\tfrac{1}{2}, x))\right\}$$

$$= \frac{8F}{a_0^2 E}\epsilon(n, x) \quad \text{by (3)}$$

$$= \frac{8F}{a_0^2 Ek}\int_{c(n)}^{x} f(n-\tfrac{1}{2}, t) \sinh k(x-t)\, dt \text{ by (15)}. \quad . \quad (16)$$

Now assume that n can be regarded as a continuous variable and that the difference eqn. (16) can be replaced by the differential equation

$$\frac{\partial f(n, x)}{\partial n} = \frac{4F}{a_0^2 Ek} \int_{c(n)}^{x} f(n, t) \sinh k(x-t)\, dt. \quad . \quad . \quad . \quad (17)$$

Assuming the work-hardening per half cycle is small, boundary condition (11) becomes

$$f(n, b(n)) = \frac{4}{Ea_0} (\Sigma_1 - \Sigma), \quad . \quad . \quad . \quad . \quad . \quad (18)$$

and (13) becomes

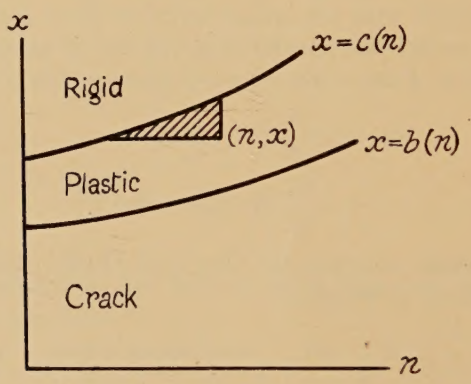
$$f(n, c(n)) = \frac{4}{Ea_0} (\Sigma_0 - \Sigma). \quad . \quad . \quad . \quad . \quad . \quad (19)$$

From (14) and (15), boundary condition (10) can be written

$$k \frac{b}{E} \Sigma = \int_{b(n)}^{c(n)} f(n, t) \cosh k(b-t)\, dt, \quad . \quad . \quad . \quad . \quad (20)$$

assuming the work-hardening per half cycle is small and that b is so large that we can take $\tanh kb = 1$.

Fig. 7



§ 6.

Integrating (17) with respect to n gives

$$f(n, x) = \frac{4}{Ea_0} (\Sigma_0 - \Sigma) + \frac{4F}{a_0^2 Ek} \iint f(u, t) \sinh k(x-t)\, dt\, du, \quad (21)$$

where the arbitrary function of integration becomes the constant $4/Ea_0(\Sigma_0 - \Sigma)$ (i.e. the value of f along the curve $x=c(n)$), provided the double integral is taken over the region in the $x-n$ plane indicated by shading in fig. 7.

Next assume that over this region of integration the boundary $x=c(n)$ can be taken as a straight line of gradient $dc/dn=c'$. Then (21) can be solved for f, either by iteration or by substituting a power series in x and n with arbitrary coefficients, giving

$$f(n, x) = \frac{4}{Ea_0} (\Sigma_0 - \Sigma) \left[1 + \frac{2F}{3a_0^2 E} \frac{(c-x)^3}{c'} + \ldots \right]. \quad . \quad . \quad (22)$$

To satisfy (18), (22) must be equivalent to

$$f(n, x) = \frac{4}{Ea_0}\left[(\Sigma_0-\Sigma)+(\Sigma_1-\Sigma_0)\left(\frac{c-x}{c-b}\right)^3\right] \quad . \; . \; (23)$$

and so

$$\frac{dc}{dn} = \frac{2F}{3a_0^2E}\,\frac{\Sigma_0-\Sigma}{\Sigma_1-\Sigma_0}\,(c-b)^3. \quad . \; . \; . \; . \; . \; . \; (24)$$

Substituting the leading term of (23) for $f(n, x)$ in (20) and taking $\cosh k(b-t) = 1$ over the range of integration gives

$$k\,\frac{b}{E}\,\Sigma = \frac{4}{Ea_0}\,(\Sigma_0-\Sigma)(c-b), \quad . \; . \; . \; . \; . \; . \; (25)$$

i.e., substituting for k,

$$(c-b) = \frac{(a_0b)^{1/2}}{2}\,\frac{\Sigma}{\Sigma_0-\Sigma}. \quad . \; . \; . \; . \; . \; . \; . \; (26)$$

From (24) and (26)

$$\frac{dc}{dn} = \frac{F}{12E}\,\frac{\Sigma^3}{(\Sigma_1-\Sigma_0)(\Sigma_0-\Sigma)^2}\,\frac{b^{3/2}}{a_0^{1/2}},$$

and from (26)

$$\frac{db}{dn} - \frac{dc}{dn} = O\left(b^{-1/2}\,\frac{db}{dn}\right),$$

therefore

$$\frac{db}{dn} = \frac{F}{12E}\,\frac{\Sigma^3}{(\Sigma_1-\Sigma_0)(\Sigma_0-\Sigma)^2}\,\frac{b^{3/2}}{a_0^{1/2}}. \quad . \; . \; . \; . \; . \; . \; (27)$$

If l is the length of the crack, i.e. $l=2b$, the integral of (27) can be written

$$l^{-1/2} = \frac{F}{24\sqrt{2}Ea_0^{1/2}}\,\frac{\Sigma^3}{(\Sigma_1-\Sigma_0)(\Sigma_0-\Sigma)^2}\,(N_\infty-n), \quad . \; . \; (28)$$

where the constant of integration N_∞ is the value of n for which the crack becomes of infinite length.

§ 7.

Equation (28) predicts that $l^{-1/2}$ should be a linear function of the number of stress cycles. Measurements of the growth of fatigue cracks have been made by Moore (1927), de Forest (1936) and Bennett (1946) and their data have been used in figs. 8–11 where $l^{-1/2}$ (with l measured in inches) is plotted against n (the origin of n being arbitrary). All three workers used rotating bending stressing which produces alternating tension compression in the test section.

The two curves of Bennett (fig. 8) are for specimens of X 4130 steel with the following tensile properties: yield stress 63 000 lb. in.$^{-2}$, ultimate tensile strength 98 000 lb. in.$^{-2}$, reduction in area at fracture 60%. The S–N fatigue curve for this material is shown in fig. 12.

De Forest (fig. 9) measured the growth of fatigue cracks in SAE 1020 steel for three different surface finishes at each of three stresses. He found that although the variation in surface finish could change the fatigue life at a given stress by a factor of four, the observed growth of the cracks

was independent of the surface state. The growth curves of fig. 9 are his superposition of measurements for three different surface finishes at each of the stress levels.

Moore (figs. 10–11) cut specimens from steel car axles which previously had been in service for a considerable period. A high fatigue stress was applied until a crack approximately 0·1 inches in length was formed. The stress was then reduced to the value shown in figs. 10–11 and the resulting growth of the crack measured. These four curves are for stresses which were less than the endurance limit of the material.

Fig. 8

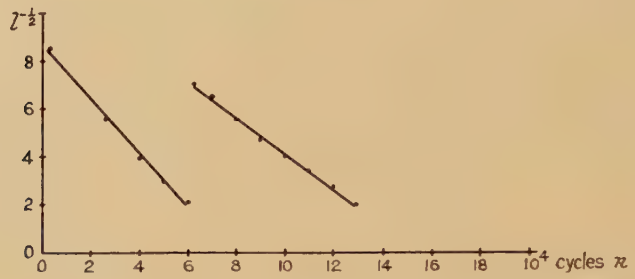
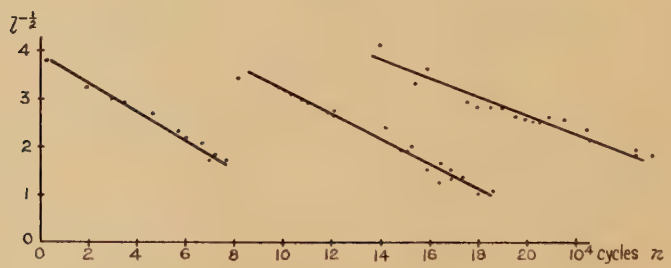


Fig. 9



In each case the experimental points show a strong tendency to lie on a straight line, the larger scatter in fig. 9 being at least partly due to the arbitrary superposition of three sets of measurements. There is also an increase in rate of growth when the crack has become large enough to affect the stress in the specimen (fig. 11). Enough data have been given by Bennett to enable also a test of the constant factor of (28). If we identify Σ_0 with the measured yield stress then the predicted ratio of the gradients of the two lines of fig. 8 is 1·51 compared with the experimental value 1·48. Thus the effect of stress level on the growth of a crack appears to be given reasonably well by the term $\Sigma^3(\Sigma_0-\Sigma)^{-2}$. In § 3 a_0 was visualized as the width of the plastic region at the tip of the crack, and a reasonable value would be $a_0=10^{-4}$ in. Taking $\Sigma_1=98\,000$ lb. in.$^{-2}$ and $E=3\cdot10^7$ lb. in.$^{-2}$ gives $F=70$ lb. in.$^{-2}$. A more convenient quantity to discuss perhaps is $(\Sigma_1-\Sigma_0)/F=500$ which, from fig. 4, is seen to be the

fracture strain of the plastic region. This value is from one hundred to one thousand times greater than the fracture strain observed in a tensile test.

There are however three effects which tend to reduce this apparent discrepancy. The fracture strain in a tensile test is governed by an instability phenomenon rather than any exhaustion of ductility in the metal, and larger strains can be obtained in other modes of deformation, e.g. torsion or rolling. Further, the actual stress–strain curve is not a straight line as assumed in fig. 4 and, as in the model of fatigue proposed by Orowan (1939), it is the plastic modulus at high strains which is important ; and this is less than F, the average value used in this model.

Fig. 10

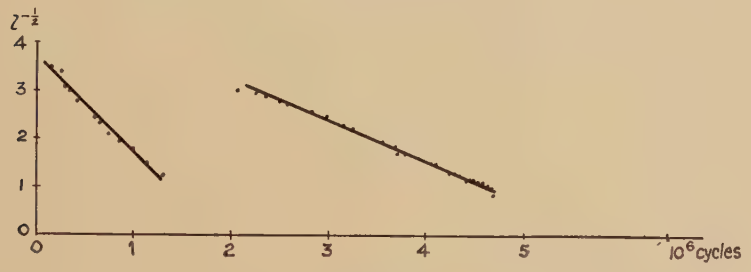
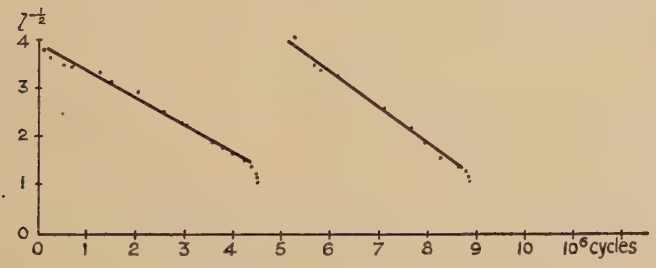


Fig. 11



Finally the Bauschinger effect has not been taken into account in the model and this would cause a large increase in ductility under alternating deformation as compared with unidirection loading. Although no quantitative argument can be given, it does not appear unreasonable that these effects could account for the above value of F.

§ 8.

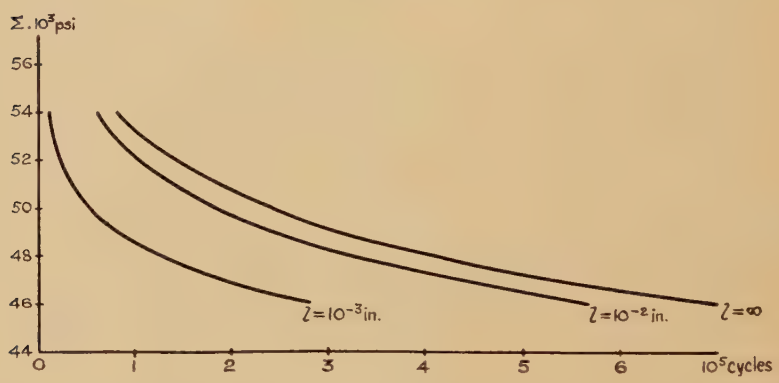
If, as figs. 8–11 suggest, $l^{-1/2}$ is a linear function of n, these straight lines can be extrapolated to find the crack length at earlier stages of the fatigue test. If this extrapolation is carried back to the start of the test, then the crack length from the two curves of fig. 8 would have been 5×10^{-4} in. and from fig. 9, 5×10^{-2} to 5×10^{-4} in. depending on the surface finish and the scatter of the experimental points. Although, as discussed

below, it is not thought that fatigue is the growth of cracks initially present, these values form a lower limit to the possible length of crack when formed. From the S–N curve given by Bennett (1946) and assuming that the effect of stress is through the factor $\Sigma^3(\Sigma_0-\Sigma)^{-2}$, contours of equal crack length can be drawn as in fig. 12. These illustrate the slow growth of small cracks and the rapid growth when large, for the contours $l=1$ in. and $l=\infty$ are effectively coincident on the scale used.

Machlin (1948) and Shanley (1952) have assumed that fatigue is simply the growth of cracks present in the initial material. There appear to be four arguments against this hypothesis.

(i) De Forest (1936) has shown that the rate of growth of cracks is independent of the surface finish of the specimen, although the total fatigue life can be changed by a factor of four. This would indicate that the surface finish influences the number of stress cycles necessary to form a crack and/or the length of crack formed.

Fig. 12



(ii) As stated in § 7 figs. 10–11 are for fatigue cracks which were produced at a high stress and then their growth at a lower stress measured. The four values of this lower stress used were all less than the endurance limit of the initial material. This suggests that fatigue involves the formation followed by growth of a crack, and that the endurance limit is the minimum stress which will form a crack rather than the minimum stress at which initially present cracks will grow.

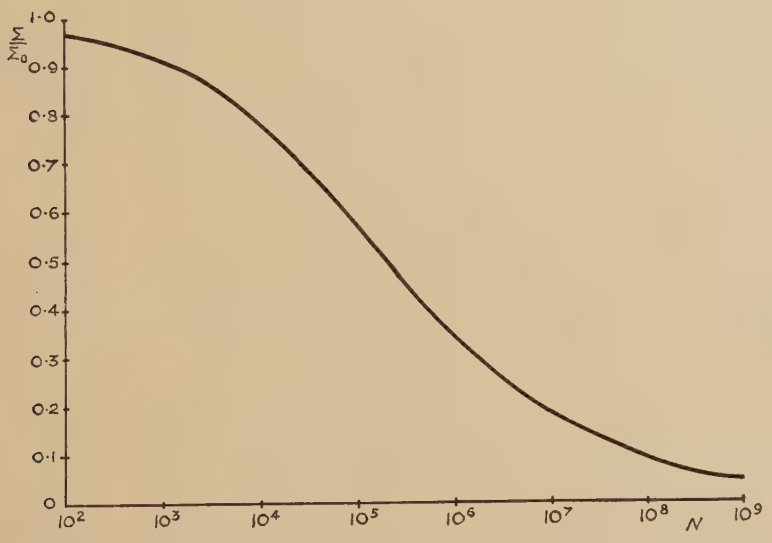
(iii) A negative argument is that there is no evidence from other properties of metals that there invariably exists cracks 10^{-4} in. long in all metals and alloys, polycrystalline or single crystal however carefully prepared.

(iv) Bennett (1946) found that the number of separate cracks visible at the end of a fatigue test varies from one at low stresses to one hundred at high stresses. If there were, say, one hundred cracks initially present in the test section of the specimen, then (28) predicts that the one which causes fracture will be the largest of these and that the length of the

others when this happens is independent of the stress and number of cycles necessary to fracture the specimen. Thus the same number of cracks would be visible at fracture at all stresses. As this is not so, it appears that the cracks are formed during the test and that more are formed at the higher stresses.

This suggests that the process of fatigue is the following: If the applied stress is above the initial yield point of the material there is bulk deformation during the first hundred cycles producing a general work-hardening. This initial stage is absent for lower stresses. During the next stage there is deformation in localized regions only, due to the inhomogeneous nature of the metal. This inhomogeneity in polycrystalline metals is probably due to the anisotropy of the crystals causing regions

Fig. 13



of higher stress, some of which may coincide with regions of less restrained slip. In single crystals there is the inhomogeneity caused by plastic flow in the first stage and the effect of any substructure present. A model of such regions has been given by Orowan (1939) where it was shown that their deformation may tend to an elastic state and the material will not fail by fatigue, or at a higher stress a crack may be formed after a finite number of cycles. There is a small amount of evidence that the crack is formed after 10–50% of the number of cycles necessary to produce failure. The remainder of the test is the growth of the crack. The model proposed here shows reasonable agreement with measurements of the later stages of crack growth. It predicts that small cracks will grow very slowly and that it is quite feasible for a crack to be growing for 90% of the fatigue life. The S–N curve as usually measured will be the sum of two parts, the number of cycles necessary to form a crack and the number of cycles

necessary for the crack to grow. For a given material the number of cycles to form a crack is a function of the stress. It seems reasonable that the size of crack formed is governed by some typical substructure dimension of the material and is relatively independent of the stress. If this is the case, then the number of cycles between creation of the crack and fracture is, by (28), of the form

$$n=C\left(1-\frac{\Sigma}{\Sigma_0}\right)^2\left(\frac{\Sigma}{\Sigma_0}\right)^{-3}, \quad . \quad . \quad . \quad . \quad . \quad (29)$$

where C is a constant depending on the initial length of the crack at creation and on the properties of the material. This expression is valid unless Σ is nearly equal to Σ_0 when the approximation used, that the work-hardening per cycle is small, is no longer true. The general form of (29) is shown in fig. 13 where C has been given the arbitrary value of 10^5. It will be seen that small variations in Σ can cause large variations of n. On combining (29) with the number of cycles necessary to produce a crack, as given by the model of Orowan (1939), realistic S–N curves may be obtained. Due, however, to the number of arbitrary parameters available in this combination, no useful purpose would be served in reproducing the result.

Acknowledgments

This work is published by permission of the Commonwealth of Australia, Department of Supply, of which the author is an Officer. I would also like to acknowledge the helpful comments of Professor N. F. Mott and Dr. N. Thompson.

References

Bennett, J. A., 1946, *Proc. Amer. Soc. Test. Mat.*, **46**, 693.
de Forest, A. V., 1936, *J. Appl. Mech.*, **53**, A, 23.
Head, A. K., 1953, *J. Mech. Phys. Solids*, **1**, 134.
Inglis, C. E., 1913, *Trans. Inst. Naval Arch.*, 219.
Machlin, E. S., 1948, *N.A.C.A. Tech. Note*, No. 1489.
Moore, H. F., 1927, *Univ. of Illinois Eng. Exp. Sta. Bull.*, No. 165.
Orowan, E., 1939, *Proc. Roy. Soc.* A, **171**, 79.
Shanley, F. R., 1952, *A Theory of Fatigue based on Unbonding during Reversed Slip* (Santa Monica, California : The Rand Corporation).
Yoffe, E. H., 1951, *Phil. Mag.*, **42**, 739.

XCIX. *A Model of the Radio-Frequency Radiation from the Galaxy*

By R. Hanbury Brown and C. Hazard
Jodrell Bank Experimental Station, Cheshire*

[Received June 23, 1953]

Abstract

A model of the Galaxy has been constructed which represents satisfactorily the radio-frequency surveys which have been made in the range 18·3 Mc/s to 1200 Mc/s. The model consists of a system of localized sources which are highly concentrated in the centre of the Galaxy and a disc of ionized interstellar gas whose properties are consistent with visual observations. It is suggested that the sources are rare members of Baade's population II and that they may be identified with the intense sources which have been observed to lie close to the galactic plane. The model requires the existence of a large isotropic component of radiation whose origin is probably extra-galactic.

§ 1. Introduction

The first theory to give a reasonable quantitative explanation of the phenomenon of extra-terrestrial radio frequency radiation was advanced by Reber in 1940. He estimated the intensity to be expected from the free–free transitions of electrons and protons in the interstellar gas, and showed that it was in rough agreement with his observations at 160 Mc/s. This theory has been discussed subsequently by a number of authors (Henyey and Keenan 1940, Townes 1947) and it has been shown that, although it explains the intensity observed at high frequencies, it fails to account for the high values of temperature† observed at low frequencies.

The discovery (Hey, Parsons and Phillips 1946, Bolton and Stanley 1948, Ryle and Smith 1948) of the localized radio sources suggested a different origin of the radiation. The relative intensity and number of the sources has been analysed by Smith (1950) and also by Bolton and Westfold (1951), and they have shown that the intensity of the radiation may be explained in terms of a system of such sources distributed widely throughout the Galaxy.

In a more detailed analysis Westerhout and Oort (1951) have compared the isophotes given by Bolton and Westfold (1950 b) at 100 Mc/s with

* Communicated by Professor A. C. B. Lovell.

† The word ' temperature ' is used in this sense throughout this paper and refers to the equivalent aerial temperature. It represents the temperature of a black-body surface which would give the observed value of received power when placed in the aerial beam.

those to be expected from the Galaxy. In constructing their model they assumed that the bulk of the radiation at 100 Mc/s arises in localized sources and they omitted the effects of the interstellar gas. Their analysis shows that reasonable agreement can be obtained between the observations and the model if it is assumed that the radio sources are distributed in the same way as the common stars. They noted, however, that in order to obtain this agreement it was necessary to assume a substantial isotropic component of radiation whose origin might be extra-galactic. Although their model shows agreement with the observations at 100 Mc/s it fails to explain the shape of the isophotes observed at either higher or lower frequencies, and it appears possible that even the agreement at 100 Mc/s may be largely fortuitous. Thus the actual isophotes used by Westerhout and Oort were obtained by a reduction of observations taken with a 17° aerial beam.* The reduction was carried out by Bolton and Westfold (1950 b) who applied the shape of the beam to the observed results and deduced the true distribution of intensity by successive approximation. This method is unreliable when the true distribution is considerably narrower than the aerial beam, and in particular, large errors may arise when the distribution is unknown in both latitude and longitude. It is therefore possible that the isophotes used by Westerhout and Oort do not represent the true distribution of radiation.

A more recent investigation by Piddington (1951) takes into account observations made over a wide range of frequencies (10 Mc/s to 1200 Mc/s). His analysis places emphasis on the spectrum of the radiation rather than on its space distribution. From the results of several observers he derives the distribution of intensity with latitude across the galactic plane and shows that it changes with frequency. He interprets this result in terms of a mixture of radio sources and ionized gas within the Galaxy, and concludes that at high frequencies ($\nu \simeq 500$ Mc/s) the ionized gas provides most of the radiation and that at low frequencies it produces significant effects by absorbing the radiation from the sources. There appear to be certain objections to these conclusions : firstly, several of the surveys on which the analysis is based were made with wide aerial beams (e.g. 17° at 100 Mc/s) and hence the derived variation of intensity with latitude cannot be regarded as reliable ; secondly, it is impossible to reconcile the shape of the isophotes observed at high frequencies with the visual evidence for the distribution of ionized gas. For example the survey made by Reber (1948) at 480 Mc/s shows a marked concentration of the isophotes in both galactic latitude and longitude and indicates that the majority of the radiation has its origin in sources which are concentrated towards the galactic centre and which lie close to the galactic plane. It may therefore be assumed that the majority of the radiation

* Throughout the present paper the value of beam-width quoted corresponds with the total width of the aerial beam between half-power points.

observed towards the galactic centre at 480 Mc/s does not arise in ionized interstellar gas, since visual observations indicate (Baade 1951) that the components of Baade's population I are rare in the nuclei of spiral nebulae similar to our own Galaxy.

The present paper presents a new attempt to interpret the available observations in terms of a simple model. The model has been designed to give agreement both with the observed values of aerial temperature and with the shape of the isophotes over the range 18·3 Mc/s to 1200 Mc/s. No attempt has been made to discuss the actual mechanism by which the radio energy is generated.

§2. A Preliminary Analysis

Previous attempts to construct a radio model of the Galaxy suggest that at least three distinct components must be taken into account: (i) localized sources within the Galaxy; (ii) ionized interstellar gas; (iii) an isotropic component which may be of extra-galactic origin.

Before developing a complete model a preliminary analysis was made to find whether the observed results could be represented without invoking the effects of ionized gas. It was assumed that all the radiation from the Galaxy arises in a system of localized sources and that the temperature observed in any direction represents the radiation from these sources together with an isotropic component. The true distribution of intensity was derived from a survey made with a narrow beam, namely that made by Reber (1948) at 480 Mc/s. Since the beam used in his survey had a width of only 4·5° the observed distribution in longitude was taken to represent the true distribution (for an infinitely narrow beam), whilst the true distribution in latitude was found by a process of trial and error. The lines of equal temperature were assumed to be confocal ellipses symmetrical about the galactic centre and their shape was found by applying the beam used by Reber to a trial set of isophotes* and by adjusting the axes of the ellipses until they reproduced the observed results. Figure 1 shows the distribution in latitude ($l'=0$) and longitude ($b=0$) observed at 480 Mc/s and the full lines show the true distributions derived from the trial isophotes. It can be seen that the true distribution in latitude is very narrow and in the direction of the galactic centre the curve has a total width between points of half intensity of about 3 degrees.

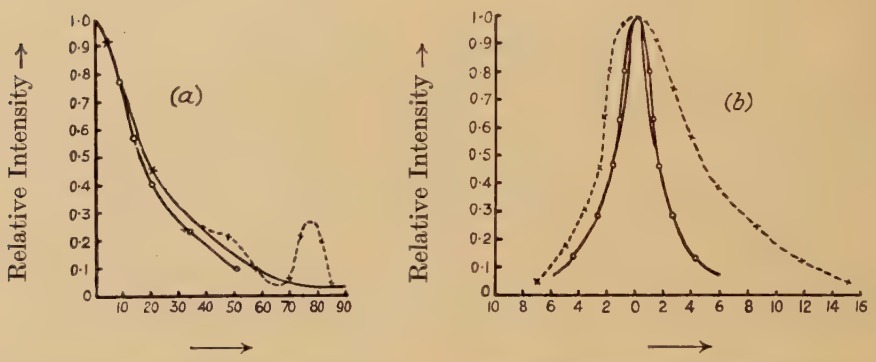
The beam shapes appropriate to the various surveys were applied to the true distributions derived from the observations at 480 Mc/s in order to predict the observed results† at 18·3 Mc/s, 100 Mc/s, 160 Mc/s, 200 Mc/s,

* The beam shape was drawn as a set of contours of equal sensitivity. It was then superimposed on the true isophotes and the observed intensity was obtained by graphical integration using a planimeter.

† This comparison has been restricted to surveys for which the actual observed results have been published. Isophotes obtained by any process of correcting the observed results for beam shapes have not been used. The authors of the surveys are shown in table 1.

and 1200 Mc/s. The resulting distributions in latitude and longitude calculated for 100 Mc/s are shown in fig. 2 together with the observed results. Following Westerhout and Oort (1951) 600°K has been subtracted from the observed results to represent the isotropic component. The agreement between the calculated and observed results is close and it was found that for all frequencies from 100 Mc/s to 1200 Mc/s the shape of the calculated and observed distributions also showed reasonable agreement. However, there is one serious discrepancy which is only apparent from a

Fig. 1



Galactic longitude (degrees) $l'(b=0)$. Galactic latitude (degrees) $b(l'=0)$.

The variation of radio intensity with galactic latitude and longitude.

– – – x – – – x – – – Observed by Reber (1948) at 480 Mc/s with a 4·5° beam.

——————— The true distributions of intensity (for an infinitely narrow beam) derived from Reber's results (see text, § 2) and adopted in the preliminary analysis to represent the radiation from the localized sources.

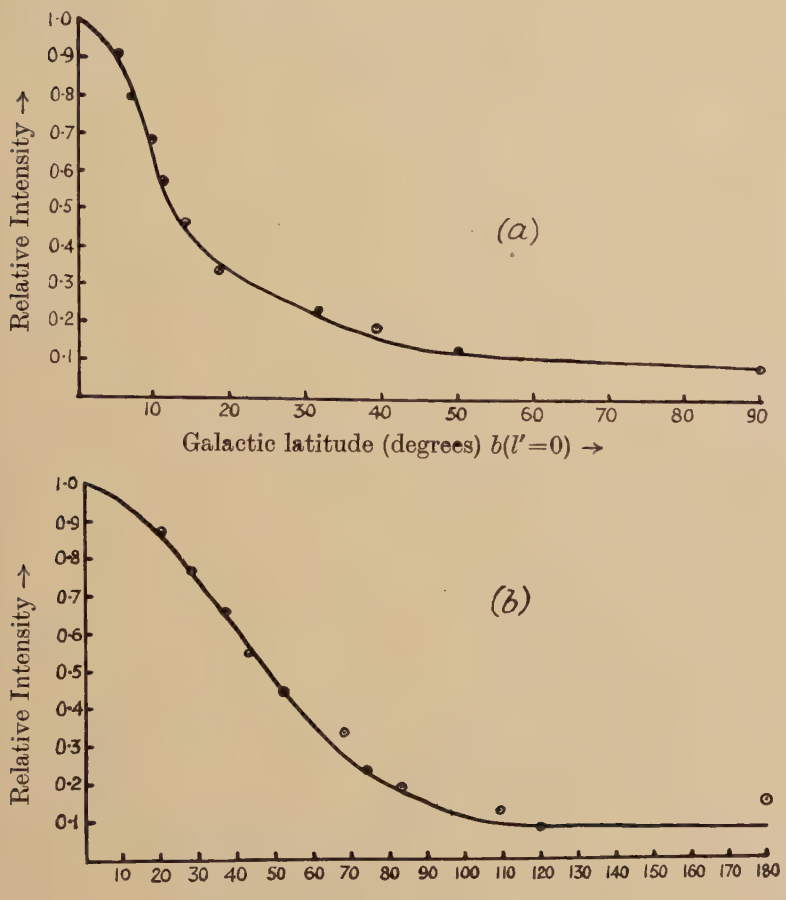
——o——o——— The true distributions of the radiation from the localized sources modified to take account of the ionized interstellar gas (see text, § 3 (b)) and adopted for the model.

detailed analysis of all the observations. This analysis shows that the observed ratio of the temperature at the galactic centre to that at the anti-centre (each temperature being measured above that at the galactic pole), decreases with increasing frequency. Thus the observed value of the ratio is about 15 at 100 Mc/s and 7 at 480 Mc/s and it is to be expected that, if the observed temperature in each survey were to be corrected for the effect of beam width, the change in ratio would be even more pronounced. This variation with frequency cannot be explained by the simple model.

A comparison between the observed results at 18·3 Mc/s and those calculated from the model reveals another significant discrepancy. Figure 3 shows some observations made by Shain (1951) with an aerial beam of approximately 40° directed to the zenith in latitude S.34°. The figure

also shows the corresponding results calculated from the model for a beam of similar shape. To facilitate comparison of the two curves an arbitrary isotropic component of 50 000°K has been included in the calculated values and the maximum temperature has been adjusted to correspond with that

Fig. 2



The variation of intensity with galactic latitude and longitude at 100 Mc/s.

———— Calculated from the true distributions shown in fig. 1 for observations with a 17° beam.

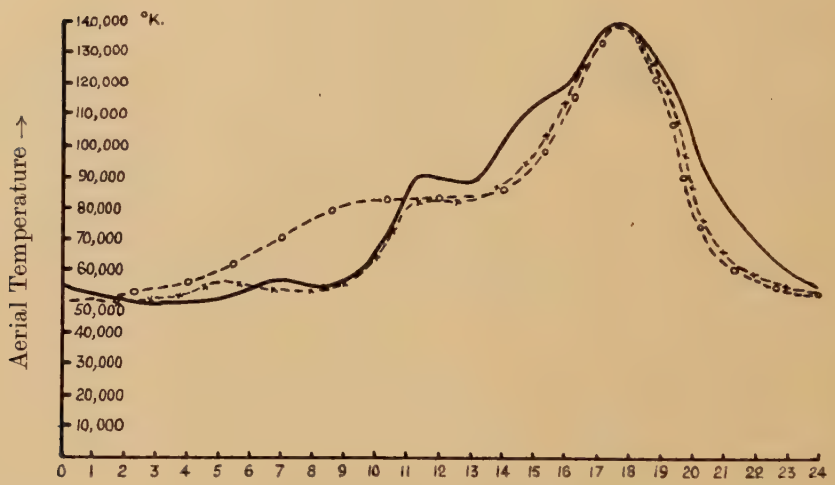
⊙ Observed by Bolton and Westfold (1950 b) using a 17° beam. For the purposes of comparison an isotropic component of 600°K (see text, § 2) has been subtracted from the observed temperature before normalizing them to unity at the galactic centre.

observed by Shain. The most marked discrepancy between the two curves occurs between R.A.5^h and 11^h, where the calculated curve shows a considerably higher intensity than the observed curve. This corresponds to a

transit of the galactic plane through the aerial beam and it is impossible to explain the minimum observed in this position by any conventional model of the distribution of sources within the Galaxy.

Thus a preliminary analysis shows that the shape of the isophotes observed towards the galactic centre at frequencies above 100 Mc/s can be represented successfully in terms of a true distribution of intensity which does not change with frequency. The distribution is more highly concentrated towards the galactic centre and into the galactic plane

Fig. 3



The variation of intensity at 18·3 Mc/s received with a beam ($\sim 40°$ wide) directed to the zenith at latitude S. 34°.

———————— Observed by Shain (1951).

– –o– – – –o– – – – Calculated from the true distributions of intensity shown in fig. 1 for the beam shape used by Shain, and neglecting the effects of ionized interstellar gas.

– – –x– – –x– – – Calculated from the model taking into account the effects of the ionized interstellar gas. The relative intensity of the radiation from the components of the model is given in table 1.

than would be expected from a source of radiation with a space-distribution similar to that given by Westerhout and Oort for the common stars. However, the simple model adopted in the analysis fails to account satisfactorily for the observed variation with frequency of the ratio of the centre to anti-centre temperatures. It also fails to account satisfactorily for the observations at 18·3 Mc/s. In the next section it will be shown that these failures can be remedied by taking into account the effects of the ionized interstellar gas.

§ 3. The Effects of the Ionized Interstellar Gas

(*a*) *A Model of the Ionized Interstellar Gas*

Consideration of the available experimental results shows that there is insufficient evidence, especially at low frequencies, to allow the properties of the ionized gas to be deduced from radio observations. Therefore in the present work it has been assumed that the gas has the characteristics deduced from visual observations and the nature of the other two components has been derived from the radio data. Following the analysis of Westerhout and Oort (1951), it has been assumed that the interstellar gas is concentrated into clouds of average diameter 10 ps with an electron density of about 5 to 10 per cm^3. It has also been assumed that a line of sight in the galactic plane crosses 5 clouds per kps* and that 10% of the clouds are ionized and have an electron temperature of 10 000°K. In their analysis Westerhout and Oort estimate that the optical depth (τ) of an average ionized cloud at 100 Mc/s is given by $\tau=0{\cdot}021$. The extension of the gas in the galactic plane is uncertain and Westerhout and Oort consider two alternative models. In the first case the gas clouds are assumed to be distributed uniformly in a cylindrical ring extending in the galactic plane from 5 to 11 kps from the centre. In the second case the clouds are assumed to have a uniform distribution over a cylindrical region extending from the galactic centre to 12·5 kps. In each case the length of the cylinder is assumed to be 200 ps. In the present work the second model of the gas has been adopted for simplicity, although the observed absence of B- and O-type stars in the nucleus of M.31 suggests that the first model may represent the visual evidence more closely.

(*b*) *The Effects of Ionized Gas on the Apparent Distribution of Radiation*

The effects of the ionized gas will be two-fold: firstly, the gas will contribute its own share to the observed radiation, and secondly, it will absorb the radiation from the isotropic component and from the sources. Before these effects can be evaluated quantitatively it is necessary to develop the equations for the radiation from a system of discrete gas clouds and localized sources. These equations are derived in the appendix and they have been used to evaluate the effects of the gas as described in the following sections.

The emission from the ionized gas

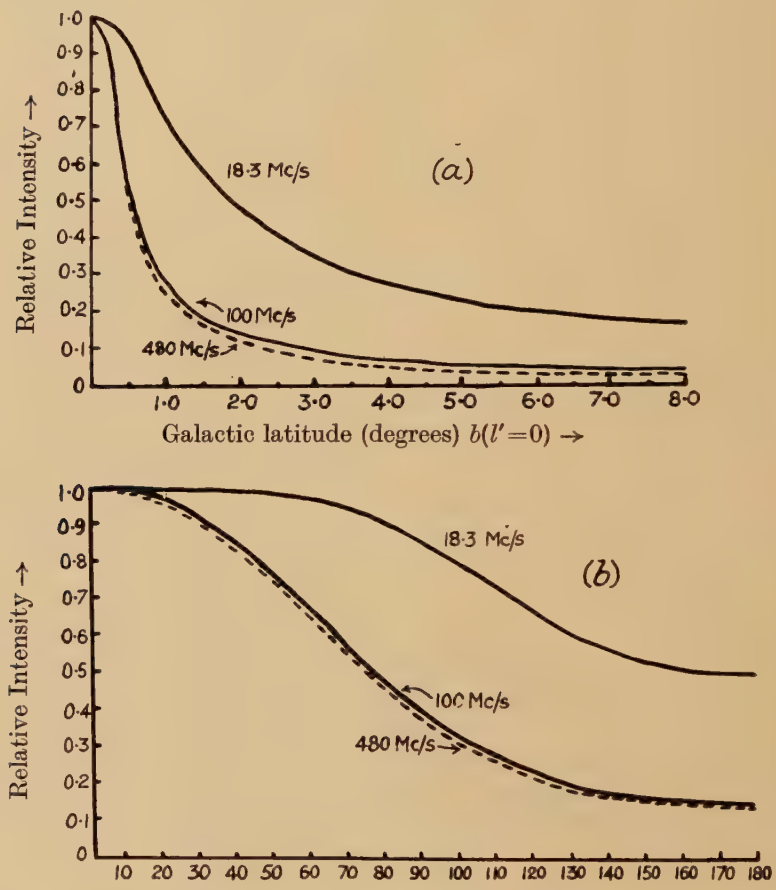
The intensity of the radiation from the ionized gas was calculated by means of eqn. (7) (Appendix) for the model of the gas given in § 3 (*a*). The distribution of this intensity with galactic latitude and longitude is shown in fig. 4 for three different frequencies. The curves represent the intensity in an infinitely narrow beam. To facilitate a comparison of their shapes they have been normalized to unity at the galactic centre and therefore do not show the variation of intensity with frequency.

* Westerhout and Oort (1951) estimate that the number of clouds lies between 5 and 10 per kps and the lower value has been arbitrarily adopted.

The distribution of intensity from the localized sources modified to take account of the ionized gas

The distribution of intensity adopted in the preliminary analysis was derived from the observations at 480 Mc/s (Reber 1948) neglecting the effects of the ionized gas, and for the complete model of the Galaxy the distribution was modified to take account of these effects. It can be shown (e.g. Smerd and Westfold 1949) that the optical depth of the

Fig. 4



The calculated distribution of the emission from the ionized interstellar gas.

ionized gas will vary approximately as ν^{-2}. Taking the optical depth of a gas cloud to be 0·021 at 100 Mc/s (see § 3 (*a*)), then at 480 Mc/s it is 9×10^{-4}. If there are 0·5 ionized clouds per kps, and the extent of the Galaxy from the sun in the direction of the galactic centre is 22 kps, then the total optical depth at 480 Mc/s is 10^{-2}. Thus absorption at 480 Mc/s is negligible and the only significant effect of the gas will be its contribution to the total radiation.

The variation of aerial beam temperature with galactic latitude and longitude due to the emission from the gas was therefore calculated for observations at 480 Mc/s with a 4·5° beam, using the distributions shown in fig. 4. These calculated variations were then used to correct* the observed beam temperatures before deriving the distribution of intensity from the sources by the method described in § 2. The modified latitude and longitude distributions are indicated in fig. 1 and they differ only slightly from those used in the preliminary analysis.

The effect of the gas absorption on the apparent distribution of radiation from the isotropic component and localized sources

The effect of absorption in the gas on the apparent distribution of intensity from the isotropic component is given by eqn. (8) (Appendix). Comparison of eqn. (8) and eqn. (7) shows that the shape of the resulting distribution is related simply to that of the gas radiation (fig. 4) and it has therefore not been shown.

The apparent distribution of the radiation from the sources in the presence of ionized gas is given by eqn. (4) (Appendix). To apply this equation it is first necessary to derive the space distribution of the sources. Taking the diameter of the Galaxy to be 25 kps and the sun to be 9·5 kps from the centre, a model was constructed in which the lines of equal source-density in the galactic plane were assumed to be circles concentric about the galactic centre. The variation of space-density in the galactic plane was then adjusted by a graphical method to fit the modified longitude distribution shown in fig. 1 (*a*). The distribution normal to the galactic plane was found by assuming that the equidensity surfaces are ellipsoids of revolution about the galactic centre, and by adjusting the ratio of their axes to give a latitude distribution which agrees with that in fig. 1 (*b*). The approximate space-distributions found in this way are shown in fig. 5. Their cosmographic significance will be discussed further in § 5 (*a*).

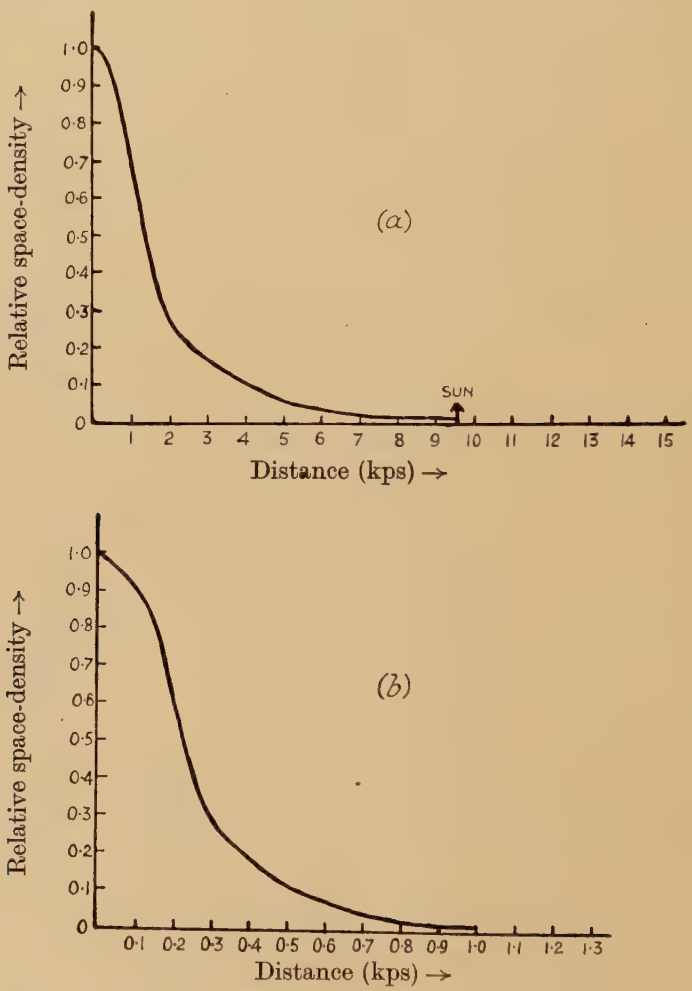
From the space-distribution of the sources and from the assumed model of the gas, the apparent distribution of the radiation from the sources can be predicted at any frequency. If the temperature observed from the sources in the presence of gas is (T_S') and the temperature which would be observed without gas is (T_S), then the ratio (T_S'/T_S) has been termed the ' gas transmission factor ' (η). From eqn. (4) (Appendix) it follows that

$$\eta = \frac{\int_0^R \rho_r \exp\left[-mr(1-\exp(-\tau))\right] dr}{\int_0^R \rho_r \, dr} \quad . \quad . \quad . \quad . \quad . \quad (1)$$

* This correction cannot be made by simple subtraction of the calculated gas radiation from the observed results since Reber's temperatures appear to be measured above a zero (or zeros) in the galactic plane. In applying the correction it has been assumed that the value given in the direction of the galactic centre represents the difference between the temperature in this direction and the point $l'=100$, $b=0$.

From eqn. (1) the value of η was evaluated graphically for several directions in space at a number of different frequencies. Using these values of η together with the distribution of the sources given in fig. 5, it was found that for all frequencies above 100 Mc/s absorption in the ionized gas has little effect on the apparent distribution of the radiation from the sources.

Fig. 5



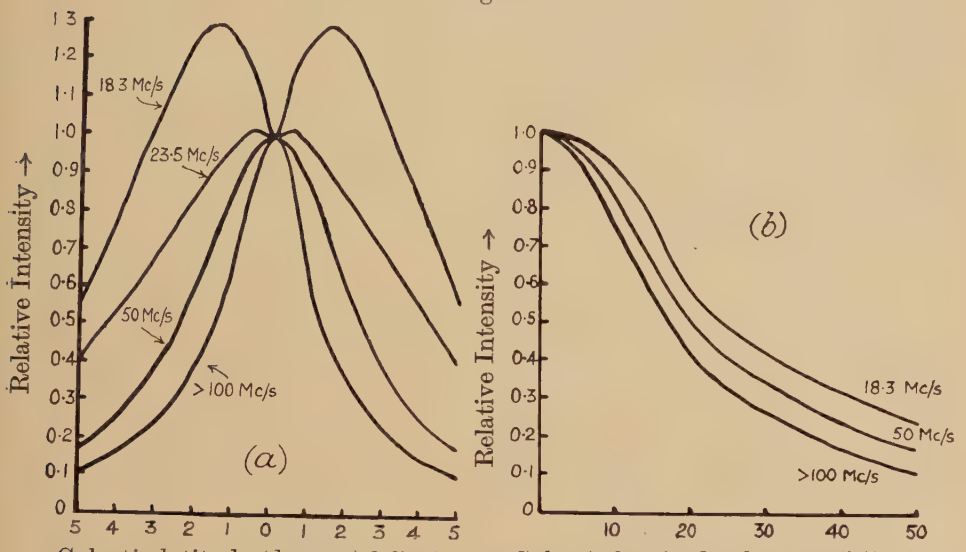
The space-distribution of the localized sources.

The curves shown have been derived from the model by the method given in the text (§ 3 (*b*)). (*a*) shows the variation of the space-density of the sources in the galactic plane with distance from the galactic centre. (*b*) shows the variation of the space-density of the sources along a line through the galactic centre normal to the galactic plane.

Below 100 Mc/s the isophotes become progressively broader and at a frequency of about 20 Mc/s the absorption in the gas is so pronounced that the appearance of a band of relatively low temperature may be

expected along the galactic plane. The distribution of intensity with latitude ($l'=0$) and longitude ($b=0$) has been calculated at various frequencies and the results are shown in fig. 6 (a). The curves refer to

Fig. 6



The apparent distribution of the radiation from the localized sources calculated from the model.

observations with an infinitely narrow beam and cannot therefore be compared directly with the results obtained with wide aerial beams.

§4. Quantitative Comparison between the Model and the Observed Results

(*a*) *Comparison in the Direction of the Galactic Centre*

Before the model could be compared with the observed results it was necessary to assign relative proportions to the three components of the radiation. This requires a knowledge of the spectra of the components together with their relative intensities at one frequency. The spectrum and the intensity of the emission from the gas has already been assumed and the parameters of the other two components were found in the following way. A separate set of isophotes was calculated for each of the three components of radiation at a number of frequencies selected to correspond with the surveys listed in table 1. The beam shapes appropriate to each survey were then superimposed on these isophotes and the effective temperature of the beam (T_B) was found by graphical integration. This temperature was expressed as a fraction of the temperature at the centre of the beam (T_C), and the ratio (T_B/T_C) was termed the ' beam reduction factor ' (f). The value of f was found for each component of the radiation with the beam centred on the galactic centre and also

Table 1. Comparison between the Observed and Calculated Beam Temperatures in the Direction of the Galactic Centre ($l'=0$, $b=0$)

	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)
1. Frequency (Mc/s)	18·3	100	160	200	480	1200	3000
2. Beam width of survey (degrees)	~40	17	12	25	4·5	2·8	3·4
3. Calculated optical depth of ionized gas ($\Sigma(\tau)=mR[1-\exp(-\tau)]$)	0·994	0·205	0·086	0·056	0·010	0·0016	0·0003
4. Gas transmission factor: *							
(a) Localized sources (η_S)	0·14	0·90	0·96	0·97	1·0	1·0	1·0
(b) Isotropic component (η_I)	0·006	0·80	0·91	0·94	0·99	1·0	1·0
5. Beam reduction factor:							
(a) Localized sources† (f_S)	0·38	0·19	0·32	0·16	0·50	0·63	0·58
(b) Ionized gas† (f_G)	0·2	0·08	0·1	0·07	0·3	0·5	0·4
(c) Isotropic component‡ (f_I)	0·8	0·98	0·99	1·0	1·0	1·0	1·0
6. Calculated temperature along a line of sight neglecting absorption in gas § (°K):							
(a) Localized sources (T_S)	$2{\cdot}1\times10^6$	18000	4850	2600	221	17	1·3
(b) Isotropic component (T_I)	35000	500	150	90	10	1	0·1
7. Calculated temperature along a line of sight taking into account gas absorption (°K):							
(a) Localized gas sources ($T_S'=\eta_S T_S$)	292000	16200	4650	2510	218	17	1·3
(b) Ionized gas ($T_G=T_e\,\Sigma(\tau)$, $T_e=10\,000$°K)	9940	2050	860	560	100	16	3
(c) Isotropic component ($T_I'=\eta_I T_I$)	210	400	136	85	10	1	0·1
8. Calculated beam temperature due to each component (°K):							
(a) Localized sources ($f_S T_S'$)	111000	3080	1490	417	109	10·7	0·8
(b) Ionized gas ($f_G T_G$)	1990	164	86	39	30	8	1·2
(c) Isotropic component ‡($f_I T_I$)	28000	490	149	90	10	1	0·1
9. Total calculated beam temperature (°K)	141000	3734	1576¶	456‖	139¶	18·7‖	2‖
10. Observed beam temperature (°K)	140000	3860**	1370	447	107	17	2·6

(i) Shain 1951. (ii) Bolton and Westfold 1950 b. (iii) Reber 1944. (iv) Allen and Gum 1950. (v) Reber 1948. (vi) and (vii) Piddington and Minnett 1951.

* The gas transmission factor in any direction for a particular component is the ratio of the actual temperature contributed by that component to the temperature which it would contribute in the absence of absorption by the ionized gas.

† The beam reduction factor for this component is the ratio of the beam temperature due to the component to the temperature which it produces along a line of sight in the direction of the galactic centre.

‡ The beam reduction factor in this case is the ratio of the beam temperature to the temperature contributed in the direction of the galactic poles.

§ As discussed in the text (§ 4 (*a*)) the spectrum of the localized sources is assumed to be $T_S \propto \nu^{-2\cdot8}$. The spectrum of the isotropic component is assumed to be $T \propto \nu^{-2\cdot5}$.

‖ The temperature given here is the sum of the source and gas contributions only. The isotropic component has been omitted since in the corresponding surveys absolute intensities were not measured but to a close approximation the observations give the difference in temperature between any point and the galactic poles. Therefore the observed temperatures do not include the isotropic component.

¶ These temperatures, which are the sum of the source and gas temperatures only, are not strictly comparable with the values in line 10, for, as explained in § 3 (*b*), the temperatures given by Reber do not represent increases relative to the galactic poles. The value given in the direction of the galactic centre appears to be the difference in temperature between the galactic centre and a point at about $l'=100$, $b=0$. The calculated differences between this point and the galactic centre at 160 Mc/s and 480 Mc/s are 1430°K and 100°K respectively which are in good agreement with the values in line 10.

** The temperatures given in the 100 Mc/s survey are aerial temperature. The beam temperature given here has been derived from the observed results by applying the correction for the aerial side lobes given by Bolton and Westfold (1950 b).

on the galactic pole. By using these values of f, together with the previously derived values of η, simultaneous equations were set up expressing the value of the observed temperature in these two directions in terms of the calculated gas radiation (T_G), the unknown temperature from the sources (T_S) and the isotropic component (T_I). The equations were solved by substituting the observed temperatures and they yielded for each frequency the values of (T_S) and (T_I). It was found that the spectrum of the localized sources can be represented closely by the law $T_S \propto \nu^{-2\cdot8}$ over the range 18·3 Mc/s to 1200 Mc/s. The spectrum of the isotropic component proved to be given by $T_I \propto \nu^{-2\cdot5}$, but must be regarded as uncertain since it is based on observations at only two frequencies (18·3 Mc/s and 100 Mc/s). The exact spectrum of this component is, however, not of great importance to the model as its effect on the shape of the distributions is only of a second order.

Using the spectra given above it was found that the observed intensity from the galactic centre at all frequencies could be represented by assuming that at 100 Mc/s the contributions from the three components of the model are :—$T_S = 18\,000°$K, $T_I = 500°$K, $T_G = 2050°$K. Table 1 shows the final values of temperature calculated from the model on these assumptions. Line 8 shows the apparent temperatures from the sources, gas, and isotropic radiation calculated for the actual beam shapes used in the surveys. The total of the three components is shown in line 9 and may be compared with the observed temperatures in line 10. It can be seen that the model represents satisfactorily the values of the temperature observed towards the galactic centre over the whole range from 18·3 Mc/s to 3000 Mc/s. The agreement in other directions will be discussed in the next section.

(*b*) *Comparison of the Observed Results with the Latitude and Longitude Distributions Calculated from the Model*

High-frequency observations. (100 Mc/s to 1200 Mc/s)

Table 1 shows that towards the galactic centre the radiation from the sources is greater than that from the gas at all frequencies between 100 Mc/s and 1200 Mc/s and line 3 shows that above 100 Mc/s the absorption in the gas can be neglected. Line 8 shows that the effects of the gas are even less important for observations made with a wide beam. In the direction of the galactic centre and for frequencies greater than 100 Mc/s, the model will therefore predict latitude distributions which are very similar to those given by the preliminary analysis in § 2, and it has already been shown that these distributions agree with the observed results. The longitude distributions predicted by the model only differ significantly from those derived in § 2 at large angles from the galactic centre where the radiation from the gas becomes more important and broadens the distribution. The temperatures at these large angles are therefore increased relative to the central temperatures compared with those given by the preliminary analysis and are in closer agreement with

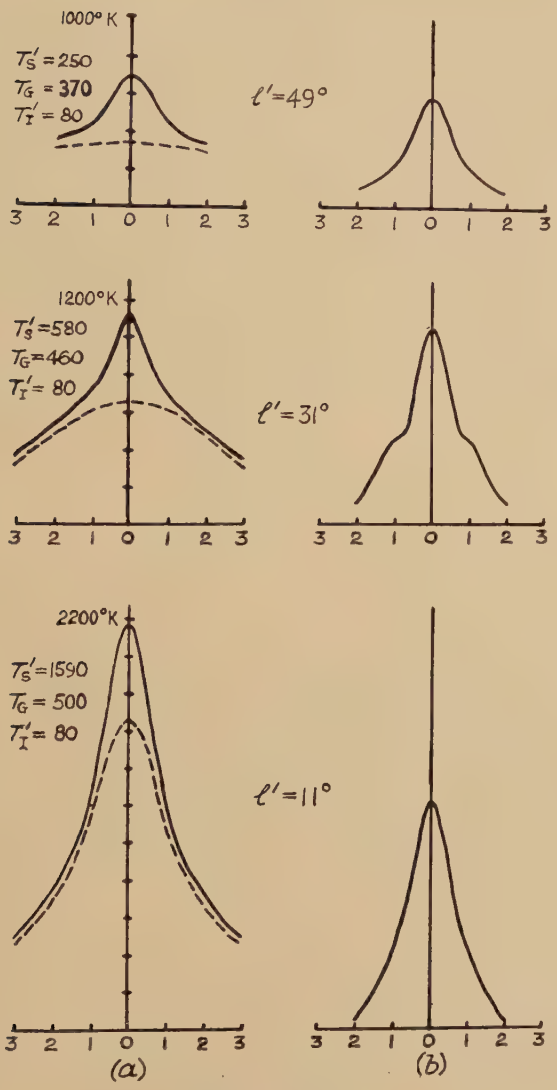
the observed results. Comparison of the longitude distributions given in figs. 4 and 6 indicates that in the direction of the anti-centre the majority of the radiation from the Galaxy arises in ionized gas. The spectrum of the radiation in the direction of the anti-centre will therefore differ from that in the direction of the centre where, for all frequencies between 1200 Mc/s and 100 Mc/s, the source radiation is predominant. It follows that over this range of frequency the ratio of the temperatures in the direction of the centre and anti-centre (each measured above the temperature at the galactic pole) will decrease with increasing frequency. The calculated change in this ratio between 480 Mc/s and 100 Mc/s is about 2·5 : 1 for an infinitely narrow beam, and about 1·6 : 1 for the beam shapes used in the surveys at these two frequencies. This calculated change of ratio is in reasonable agreement with the observed value of 15 : 7 quoted in § 2. The introduction of the ionized gas therefore removes one of the discrepancies shown by the simple model of the Galaxy used in the preliminary analysis. For frequencies much greater than 1200 Mc/s the gas radiation will become predominant in the direction of the centre and the ratio of the centre and anti-centre temperatures may be expected to tend to a constant value.

It must, however, be noted that the model predicts a smooth decrease of intensity in longitude from the galactic centre to the anti-centre, whereas inspection of the available surveys shows that the actual distribution is irregular. For example, there are pronounced maxima at $l=40°$ and $l=240°$ and a shallow minimum at $l=200°$. It seems probable that these irregularities are associated with certain features of the structure of the Galaxy which have not been taken into account. Bolton and Westfold (1950 a) have attributed the maximum at $l=40°$ to the radiation from a spiral arm and it is possible that the maximum at $l=240°$ and the minimum at $l=200°$ are also associated with the spiral system.

Low-frequency observations. (<100 Mc/s)

At low frequencies absorption by the ionized gas has a marked effect on the apparent distribution of the radiation from the sources and from the extra-galactic component. The most suitable low-frequency survey for comparison with the model is that made by Shain (1951) at 18·3 Mc/s and fig. 8 shows a set of isophotes of the apparent distribution of the radiation from the sources calculated for this frequency. The isophotes show clearly that the model predicts the presence of a band of low intensity along the galactic plane. The broken line crossing the isophotes represents the path followed by the centre of the beam used by Shain in his survey at 18·3 Mc/s. The full line in fig. 3 shows the observed variation of aerial temperature and the broken line shows the corresponding values calculated from the model. The calculated values were obtained by first passing the beam used by Shain over the isophotes in fig. 8, thus finding the contribution due to the sources alone. The contribution from the ionized gas and the isotropic component were then found by a

Fig. 7



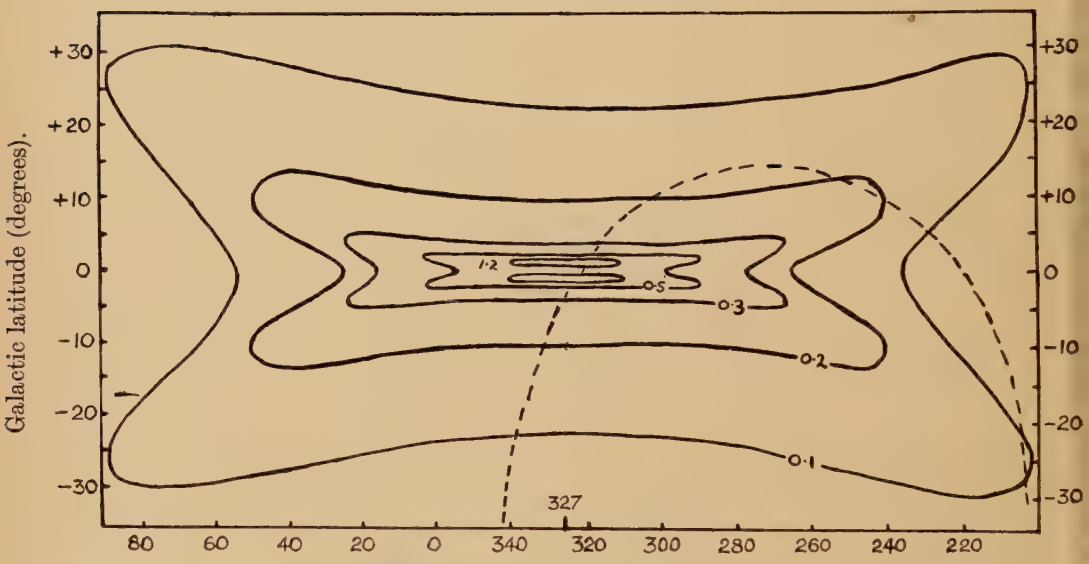
Galactic latitude

The distribution of intensity with galactic latitude at 210 Mc/s.

The curves on the left show the variation of aerial beam temperature (for an infinitely narrow beam) for 210 Mc/s calculated from the model. The full line shows the total temperature due to all three components and the broken line shows that due to the sources and isotropic component only. The relative temperature due to the sources (T_S'), gas (T_G) and the isotropic component (T_I') are shown for each value of longitude at the position ($b=0$). The curves on the right show the values of *relative* temperature observed by Scheuer and Ryle. They have been plotted to the same scale as the calculated curves and the ordinate scale of absolute temperature has been omitted.

similar analysis and the temperatures due to the three components were then added to give the calculated values of total aerial temperature. The agreement between the calculated and the observed curves is more satisfactory than that obtained in the preliminary analysis which neglected the effects of the ionized gas. The minimum at 09^h ($l'=225°$), which previously proved difficult to explain (see § 2), is predicted correctly by the model and inspection of fig. 8 shows that it is a consequence of absorption by the gas in the galactic plane. Other features of the curve, such as the plateau from R.A. 11^h ($l'=245°$) to R.A. 13^h ($l'=275°$) also appear to be due to absorption. It is possible that the differences between R.A. 20^h and 23^h and at R.A. 15^h are due to irregularities in the galactic structure, or to the presence of individual localized sources which have not been considered. For example, Shain notes that the intense source in Centaurus may be responsible for the increase of intensity observed at R.A. 15^h.

Fig. 8



Isophotes of the radiation from the localized sources in the Galaxy calculated for 18·3 Mc/s.

The isophotes show the variation of intensity from the sources alone calculated from the model taking into account absorption in the interstellar gas. They represent the relative intensity received by an infinitely narrow beam and have been normalized to unity at the galactic centre. The broken line shows the path of the beam used by Shain (1951).

Some recent results of Scheuer and Ryle

Since the construction of the model described in the present paper the results of some new measurements by Scheuer and Ryle have been

received by the authors,* and it is of interest to compare them with the predictions of the model. Scheuer and Ryle using an interferometer, have measured the apparent distribution of intensity across the galactic plane at frequencies of 81·5 Mc/s and 210 Mc/s.

The 81·5 Mc/s observations were made at two values of galactic longitude corresponding to $l'=26°$ and $l'=46°$. The temperatures measured in the galactic plane ($b=0$) at these two values of l' are given in table 2 where they are compared with the values calculated from the model.

Table 2. Comparison of the Temperatures in the Galactic Plane ($b=0$) as Observed by Scheuer and Ryle at 81·5 Mc/s with those Calculated from the model

Galactic longitude degrees (l')	Calculated temp. from sources	Calculated gas temp.	Calculated isotropic temp.	Total calculated temp.	Temp. observed by Scheuer and Ryle
26	9000	2900	600	12500	12400
46	3750	2500	630	6880	9400

It can be seen that the agreement at ($l'=26$, $b=0$) is very close but there is a considerable discrepancy at ($l'=46$). This discrepancy is not considered to be serious, for not only is the source distribution rather uncertain at large angles from the galactic centre, but an examination of the 480 Mc/s, 200 Mc/s and 100 Mc/s contours shows that near ($l'=40$) there is an irregularity in the background radiation. It is possible that this irregularity accounts for the difference between the calculated and observed temperatures. The shape of the distributions of intensity with latitude observed at both ($l'=26°$) and ($l'=46°$) have also been compared with the model and show close agreement.

In the experiments at 210 Mc/s absolute values of aerial temperature were not determined and a quantitative comparison with the model cannot be made. At this frequency the change in temperature with latitude near the galactic plane was measured at three values of galactic longitude. The results are shown in fig. 7 (*b*) and for comparison the corresponding distributions of absolute temperature calculated from the models are shown in fig. 7 (*a*).

The most interesting feature of the observed curves is the change of shape with longitude. This change of shape is reproduced by the calculated curves and appears to be due to the different longitude distributions of the sources and the ionized gas (see figs. 4 and 6). Thus at ($l'=11°$) the gas radiation is relatively weak and the observed distribution

* The authors are indebted to Mr. M. Ryle for making these measurements available prior to their publication (see *Mon. Not. Roy. Astr. Soc.*, **113**, 3, 1953).

corresponds closely to that of the sources alone. At ($l'=31°$) the contributions from the gas and the sources are approximately equal and the narrow latitude distribution of the gas can be seen superimposed on the broader distribution of the sources. At ($l'=49°$) the gas radiation is predominant and almost entirely determines the shape of the curve. This general agreement between the observed and calculated results gives support to the distributions adopted in the model. A detailed analysis shows that an even closer agreement can be obtained if the value assumed for the intensity of the gas component is increased by about 50%. The gas contribution evaluated in § 3 (*a*) is based on the assumption that a line of sight in the galactic plane intersects 0·5 ionized clouds per kps and the necessary increase can be obtained by increasing this value to 0·75 kps. This increased density is still consistent with the optical evidence and in fact lies between the limits estimated by Westerhout and Oort.

§ 5. Discussion

The analysis presented in the previous sections shows that the high frequency observations in the direction of the galactic centre may be explained successfully by assuming that the bulk of the radiation in this direction arises in a group of sources which are highly concentrated in the galactic plane and towards the galactic centre. The observed change in shape of the isophotes at frequencies above 100 Mc/s is primarily due to the different beam widths used in the various surveys. The observed difference in spectrum between the galactic centre and anti-centre, together with the distribution of intensity observed at frequencies below 100 Mc/s, may be explained by taking into account the effects of the ionized interstellar gas whose properties are known from visual observations. It is also necessary to assume an isotropic component of radiation whose origin may be extra-galactic. The model presented must necessarily be crude since most of the experimental data has been obtained with instruments of low resolving power ; furthermore, very few observations have been made at low frequencies. Nevertheless the agreement between the model and the observed results suggests that its main features are correct and the implications of the model will now be examined in further detail.

(*a*) *The Localized Sources*

It has been assumed in constructing the model that the Galaxy contains a population of localized sources whose spectra are all identical. It is of interest to test if the properties derived for these hypothetical sources are consistent with the sources which have actually been observed. If α=absolute intensity of an average source in watts/steradian/c.p.s, ρ_0=space-density of sources in the centre of the Galaxy in sources/cu. ps, ρ_r=space-density of sources at a distance r parsecs from the sun, then the temperature (T_S) due to the sources in any direction will be

$$T_S = \frac{\alpha \rho_0 c^2}{2\kappa\nu^2} \int_0^R \frac{\rho_r}{\rho_0}\, dr, \qquad \ldots\ldots\ldots \quad (2)$$

where R ps is the extent of the Galaxy in that direction and κ is Boltzmann's constant. The value of $\int_0^R \rho_r/\rho_C \, dr$ was found for the direction of the galactic centre by graphical integration of the distribution shown in fig. 5 and was substituted in eqn. (2) together with $T_S = 18\,000°$K, and $\nu = 100$ Mc/s. The value of $\alpha\rho_C$ at 100 Mc/s was found to be 1×10^{10} watts/steradian/c.p.s./cu. ps. If (ρ_S) is the density of sources near the sun then fig. 6 shows that $\rho_S \simeq 0{\cdot}01\rho_C$ and hence the power emitted by the sources in a cubic parsec near the sun is given by

$$\alpha\rho_S \simeq 10^8 \text{ watts/steradian/c.p.s./cu. ps.}$$

It has been shown by Mills (1952) and independently by Hanbury Brown and Hazard (1953) that the observed sources can be divided into two classes. The most intense sources (class I) show a marked concentration into the galactic plane, while the weaker sources (class II) show a more isotropic distribution. A rough analysis of the class I sources between $l = 40°$ and $l = 130°$ has been given by the authors (Hanbury Brown and Hazard 1953) on the simple assumption that there is no dispersion in their absolute intensities. It shows that the observed sources may be represented as members of a population in the Galaxy with a space-density near the sun of $\rho_S \simeq 5 \times 10^{-8}$ sources/cu. ps and an absolute intensity $\alpha \simeq 10^{15}$ watts/steradian/c.p.s. at 158·5 Mc/s. Assuming that the corresponding value of α at 100 Mc/s is $1{\cdot}6 \times 10^{15}$ watts/steradian/c.p.s. then the total power emitted by these sources near the sun is 8×10^7 watts/steradian/c.p.s./cu. ps. These values agree reasonably well with those deduced by Mills from independent data. The total radiation per cubic parsec from the class I sources near the sun is therefore roughly equal to the value found for the sources in the model ; and it will be assumed tentatively that the class I sources can be identified* with the population in the model. Since the radiation from the Galaxy can be explained in terms of class I alone it is unnecessary to invoke a contribution from the sources of class II and the possibility remains that the majority of the latter are extra-galactic.

The model shows that the majority of the sources within the Galaxy must be concentrated into an ellipsoidal volume in the centre. The major axis of this volume lies in the galactic plane and has a length of about 6 ps ; the minor axis is normal to the plane, and has a length of about 1 ps. The dimensions of the minor axis must be regarded as uncertain since the distributions of intensity with latitude have been derived from observations made with aerial beams of width large compared with the true distributions. Outside this central nucleus the space-density of the sources falls off slowly and reaches a low value in the neighbourhood of the sun.

* The spectrum derived for the sources in the model cannot be used to check this identification since no measurements of the spectrum of a class I source are yet available.

The distribution of the sources implies that they are members of Baade's population II.* Their space-density is uncertain and cannot be derived accurately without a knowledge of the dispersion of their absolute intensities. The rough value given above ($\rho_S \simeq 5\times10^{-8}$ sources/cu. ps) suggests that they must be rare objects with a population density about one million times less than that of the common stars. The space-density and distribution of the localized sources resembles that of some rare members of population II and may be compared, for example, with that of the planetary nebulae (Minkowski 1951).

The space-distribution of the sources used in the present model differs significantly from that given by Bolton and Westfold (1951). The model shows that the sources have a greater concentration towards the centre of the Galaxy than that given by their work. Bolton and Westfold derived their distribution from isophotes reduced from their observations at 100 Mc/s and the analysis presented in § 2 suggests that these isophotes are considerably broader than the true distribution.

The space-density suggested for the sources represents a much lower value than has previously been deduced. Thus Smith (1950) has derived a value of 3 sources/cu. ps, Bolton and Westfold (1951) give 2×10^{-4} sources/cu. ps, and Westerhout and Oort (1951) find a value of $3{\cdot}2\times10^{-4}$ sources/cu. ps, for a population of sources with no dispersion in absolute intensity. These values are all based on the assumption that the class II sources are members of the Galaxy which may be expected to lead to a much higher density than that assumed in the present work.

The suggestion that the majority of the radiation at 100 Mc/s arises from a population of class I sources disagrees with a statement by Mills (1952) who noted that the integrated output from the class I sources is unlikely to account for the background radiation from the Galaxy. Mills assumed that the density of class I sources throughout the Galaxy is uniform and equal to the observed density near the sun. The present analysis indicates that this assumption is incorrect and that the density of sources in the centre of the Galaxy is much greater than that near the sun.

(*b*) *The Ionized Interstellar Gas*

The analysis presented in §§ 3 and 4 shows that all the available radio observations can be interpreted in terms of a simple model of the Galaxy which contains a disc of ionized interstellar gas. The properties of this gas necessary to explain the observations are consistent with those derived from visual observations. The available experimental results especially at low frequencies are not adequate to allow the detailed

* Although it has been assumed that the space-distribution of the sources given in fig. 5 implies that they are true members of Baade's population II, the possibility cannot be excluded that the localized radio sources arise from the interaction of objects of population II with the interstellar gas. This question might be decided if satisfactory observations could be made of an assembly of pure population II, for example from a suitable spheroidal galaxy.

properties of the gas to be deduced. It does appear, however, that the total opacity of the gas adopted in the model is rather low and that better agreement with observation could be obtained by assuming that a line of sight in the galactic plane intersects 0·75 clouds per kilo-parsec and that an average cloud has an opacity of 0·021 at 100 Mc/s and an electron temperature 10 000°K.

(*c*) *The Isotropic Component*

The model assumes the existence of a large isotropic component which accounts for about 60% of the total radiation observed at 100 Mc/s. The magnitude of this component must be regarded as uncertain since it depends critically on the measurement of the temperature of the galactic poles and this measurement is particularly liable to experimental errors.

From the available evidence it is reasonable to assume that the origin of this component of radiation is extra-galactic. A number of the weak sources (class II) have been found to correspond in position with extra-galactic nebulae, furthermore the present analysis shows that the radiation from the Galaxy may be explained solely in terms of a population of class I sources. It is thus possible that the majority of the class II sources are extra-galactic and the isotropic component may represent the total effect of their radiation. This possibility will be examined in detail in a separate paper.

(*d*) *The Spectrum of the Radiation*

Section 4 shows that the variation with frequency of the temperature due to the three components of the model may be represented as follows : temperature from sources, $T_S \propto \nu^{-2\cdot8}$; temperature from isotropic component, $T_I \propto \nu^{-2\cdot5}$; temperature from ionized gas, $T_G \propto \nu^{-n}$; where $n=2$ for frequencies greater than about 100 Mc/s and $0<n<2$ for lower frequencies. The spectrum of the isotropic component is based on measurements at only two frequencies (18·3 Mc/s and 100 Mc/s) and is therefore uncertain.

Since both the exponent of ν and the space-distribution differ for each component of the radiation it follows that the exponent of ν for the total radiation will vary both with frequency and with the direction of observation. For example the temperature observed towards the galactic centre will vary with frequency in a complex manner. At a frequency of about 100 Mc/s the sources and the isotropic component will contribute the majority of the radiation and the exponent of ν will lie between $-2\cdot8$ and $-2\cdot5$. At higher frequencies the emission from the gas will increase relative to the other two components and the exponent of ν will therefore decrease with increasing frequency, tending to a value of 2·0 at extremely high frequencies. ($\nu \simeq 3000$ Mc/s.) At frequencies between 100 Mc/s and 18·3 Mc/s the gas clouds will become increasingly opaque

and although the majority of the radiation will arise in localized sources the exponent of ν will be less than 2·8 due to absorption in the gas. The spectrum below 18·3 Mc/s cannot be forecast from the model since the spectrum of the sources cannot be extrapolated to lower frequencies without further experimental evidence.

It has already been shown in § 4 (*a*) that the spectra assumed in the model give satisfactory agreement with the values of temperature observed over a wide range of frequencies. They may also be compared with the spectra derived by other authors. Thus Moxon (1946) has shown that in the range 40 Mc/s to 200 Mc/s the temperature of an aerial directed towards a region near the galactic centre varies as $\nu^{-2\cdot69}$, and that for regions away from the galactic plane the temperature varies as $\nu^{-2\cdot1}$. Stanley and Slee (1950) have shown that over approximately the same range of frequencies the temperature near the galactic centre varies as $\nu^{-2\cdot65}$ and near the poles as $\nu^{-2\cdot35}$. These observations confirm that the exponent of ν is higher towards the galactic centre than towards the poles. A more detailed analysis of the frequency spectrum in certain selected regions has been made by Piddington (1951). He found that in the direction of the galactic centre the observed exponent of ν varies with frequency, being about $-2\cdot0$ for frequencies greater than 1000 Mc/s, $-2\cdot51$ for frequencies around 100 Mc/s and lying between 0 and -2 for frequencies less than 40 Mc/s. This spectrum shows rough qualitative agreement with the present model ; however, the quantitative agreement is not satisfactory. The discrepancies cannot be regarded as a serious objection to the model since the distribution of intensity towards the galactic centre is very narrow, and in such a region the intensities deduced from surveys made with wide aerial beams will depend to a large extent on the method adopted for correcting for the effects of the aerial beam.

(*e*) *The Total Radiation from the Galaxy*

The total emission from the sources in the Galaxy (α_G) was calculated by integrating the space-distributions derived in § 3 (*b*) and was found to be $3\cdot2\times10^{10}\alpha\rho_C$. Substituting the value $\alpha\rho_C=10^{10}$ watts/steradian/c.p.s./cu. ps derived in § 5 (*a*) it follows that the total emission from all the sources in the Galaxy is $3\cdot2\times10^{20}$ watts/steradian/c.p.s.* at 100 Mc/s.

The total emission from the ionized gas in the Galaxy has been calculated to be $1\cdot5\times10^{19}$ watts/steradian/c.p.s. at 100 Mc/s. Thus at 100 Mc/s where the effects of absorption in the gas can be neglected, the total radiation from the model is $3\cdot4\times10^{20}$ watts/steradian/c.p.s.† where 95%

* If the value of α for an individual source is taken to be $1\cdot6\times10^{15}$ watts/steradian c.p.s. at 100 Mc s as suggested in § 5 (*a*), this intensity implies that the total number of sources in the Galaxy is about 10^5.

† The accuracy of the construction of the model has been checked by direct integration of the observed isophotes at 100 Mc s. This integration shows that, neglecting the isotropic component, the total radiation from the Galaxy is about 5×10^{20} watts/steradian/c.p.s., which is good agreement with that calculated from the model.

of the radiation arises in localized sources and 5% is due to ionized gas. At frequencies higher than 100 Mc/s, due to the different spectra of the gas and the localized sources, the proportion of the radiation due to the ionized gas will increase until at extremely high frequencies (>3000 Mc/s) the majority of the radiation will arise in the gas. At frequencies less than 100 Mc/s the apparent division of the radiation between the two components will be complicated by the effects of absorption in the gas, and it is interesting to note that at low frequencies the intensity from a galaxy containing interstellar gas should depend on its orientation with respect to the observer.

Acknowledgments

The work has been carried out at the Jodrell Bank Experimental Station of the University of Manchester. We wish to thank Professor A. C. B. Lovell for his interest in the investigation. One of us (R. Hanbury Brown) is indebted to I.C.I. Limited for a research fellowship.

APPENDIX

The Radiation from a Mixture of Ionized Gas Clouds and Localized Sources

The effective temperature observed in any direction is assumed to be due to the radiation from ionized gas clouds, together with the radiation from the sources and the isotropic component both reduced by absorption in the gas. It is assumed that the isotropic component is extra-galactic and that it must therefore traverse all the ionized gas in any direction. The effect of the gas on the intensity observed from the sources will be considered first.

Let a line of sight from the sun in any direction intersect an average number (m) gas clouds per kps, and let the average optical depth of a cloud be (τ) and the electron temperature be T_e. Consider an aerial beam of solid angle (ω) to receive radiation from sources lying in an elementary volume between r and $r+dr$ and let the number of sources in this interval of distance be N and the intensity received from a single source in the absence of gas clouds be I. For a random array of clouds the probability (ϕ_n) that there are (n) clouds traversed in any given direction is given by

$$\phi_n = \frac{(mr)^n}{n\,!} \exp(-mr). \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

Therefore the number of sources between (r) and ($r+dr$) which are obscured by (n) clouds is $N\phi_n$. The total intensity from these sources (dI') is

$$dI' = I \exp(-n\tau) N\phi_n.$$

The total intensity (I'_{dr}) from all the sources in the elementary volume will be

$$I'_{dr}=\sum_{n=0}^{\infty} I\exp(-n\tau)N\phi_n.$$

Substituting for ϕ_n from eqn. (1),

$$\left.\begin{aligned} I'_{dr}&=NI\exp(-mr)\sum_{n=0}^{\infty}\frac{(mr)^n}{n!}\exp(-n\tau)\\ \therefore I'_{dr}&=NI\exp\{-mr(1-\exp(-\tau))\}.\end{aligned}\right\} \quad . \ . \ . \quad (2)$$

If the density of the sources at a distance r is ρ_r and the absolute intensity of an average source is α watts/steradian/c.p.s. then eqn. (2) may be written

$$I'_{dr}=\alpha\omega\rho_r\exp\{-mr(1-\exp(-\tau))\}\,dr,$$

and the total intensity from all sources within a distance R will be

$$I'=\alpha\omega\int_0^R\rho_r\exp\{-mr(1-\exp(-\tau))\}\,dr. \quad . \ . \ . \ . \quad (3)$$

From the Rayleigh-Jeans formula the effective aerial temperature (T_S') observed in this direction due to sources will be given by*

$$T_S'=\frac{\alpha c^2}{2\kappa\nu^2}\int_0^R\rho_r\exp\{-mr(1-\exp(-\tau))\}\,dr. \quad . \ . \ . \quad (4)$$

Equation (4) predicts that when the gas clouds become opaque ($\tau\gg 1$) the temperature due to the sources does not fall to zero but tends to a value given by

$$T_S'=\frac{T_S\int_0^R\rho_r\exp(-mr)\,dr}{\int_0^R\rho_r\,dr} \quad . \ . \ . \ . \ . \ . \quad (5)$$

where T_S is the temperature from the sources in the absence of gas clouds. This residual temperature represents the radiation from the nearby sources and is a consequence of the model assumed.

The radiation due to the gas clouds may be derived as follows. In any direction in which the extent of the gas is R the probability (ϕ_n) of intersecting n clouds in the line of sight is

$$\phi_n=\frac{(mR)^n}{n!}\exp(-mR), \quad . \ . \ . \ . \ . \ . \quad (6)$$

and the effective temperature (T_g) in this direction will be

$$T_g=T_e(1-\exp(-n\tau)).$$

---

* Dr. J. H. Piddington has made a previous analysis of the radiation from a mixture of sources and gas and the authors wish to acknowledge the benefit of his unpublished work. He treats the case where the density of sources is uniform throughout the gas, and it can be shown that eqn. (4) above reduces to his result in this special case.

The average temperature (T_G) observed over a solid angle large compared with one cloud will be

$$\left.\begin{aligned} T_G &= \sum_{n=0}^{\infty} T_e[1-\exp(-n\tau)]\frac{(mR)^n}{n!}\exp(-mR) \\ \therefore T_G &= T_e\{1-\exp[-mR(1-\exp(-\tau))]\}. \end{aligned}\right\} \quad . \quad . \quad (7)$$

It follows from eqn. (7) that if T_I represents the temperature in any direction due to the isotropic component in the absence of absorption by gas, then the observed isotropic temperature T_I' will be given by

$$T_I' = T_I \exp\{-mR(1-\exp(-\tau))\}. \quad . \quad . \quad . \quad . \quad (8)$$

From eqns. (4) and (8) the effects of the ionized gas on the sources and on the isotropic component can be evaluated, and from eqn. (7) the radiation from the gas itself can be found.

References

Allen, C. W., and Gum, C. S., 1950, *Austr. J. Sci. Res.* (A), **3,** 224.

Baade, W., 1951, *Publications of the Observatory of the University of Michigan,* **10,** 7.

Bolton, J. G., and Stanley, G. J., 1948, *Nature, Lond.*, **161,** 312.

Bolton, J. G., and Westfold, K. C., 1950 a, *Nature, Lond.*, **165,** 487 ; 1950 b, *Austr. J. Sci. Res.* (A), **3,** 19 ; 1951, *Ibid.*, (A), **4,** 476.

Hanbury Brown, R., and Hazard, C., 1953, *Mon. Not. Roy. Astr. Soc.*, **113,** 123.

Henyey, L. G., and Keenan, P. C., 1940, *Astrophys. J.*, **91,** 625.

Hey, J. S., Parsons, S. J., and Phillips, J. W., 1946, *Nature, Lond.*, **158,** 234.

Mills, B. Y., 1952, *Austr. J. Sci. Res.* (A), **5,** 266.

Minkowski, R., 1951, *Publications of the Observatory of the University of Michigan,* **10,** 25.

Moxon, L. A., 1946, *Nature, Lond.*, **158,** 758.

Piddington, J. H., 1951, *Mon. Not. Roy. Astr. Soc.*, **111,** 45.

Piddington, J. H., and Minnett, H. C., 1951, *Austr. J. Sci. Res.* (A), **4,** 459.

Reber, G., 1940, *Proc. I.R.E.*, **28,** 68 ; 1944, *Astrophys. J.*, **100,** 279 ; 1948, *Proc. I.R.E.*, **36,** 1215.

Ryle, M., and Smith, F. G., 1948, *Nature, Lond.*, **161,** 312.

Shain, C. A., 1951, *Austr. J. Sci. Res.* (A), **4,** 258.

Smerd, S. F., and Westfold, K. C., 1949, *Phil. Mag.*, **40,** 831.

Smith, F. G., Reported by Ryle, M., 1950, *Rep. Prog. Phys.*, **13** (London : Physical Society), p. 184.

Stanley, G. J., and Slee, O. B., 1950, *Austr. J. Sci. Res.* (A), **3,** 234.

Townes, C. H., 1947, *Astrophys. J.*, **105,** 235.

Westerhout, G., and Oort, J. H., 1951, *B.A.N.*, **11,** 323.

C. *Oscillator Switching for Variable Frequency Synchrotron Control*

By D. E. Caro* and L. U. Hibbard
Department of Physics, University of Birmingham†

[Received June 19, 1953]

Abstract

A method is discussed for switching from one sweeping oscillator to another or to a fixed oscillator, without significant discontinuity of frequency, phase or amplitude.

The particle acceleration frequency in the Birmingham proton synchrotron (Oliphant 1947, Hibbard 1950) is derived from a mechanically tuned beat oscillator (Caro and Hibbard 1952). A circuit based on the above method is described, by means of which the output frequency is locked at a specified value when reached by the sweeping oscillator.

§ 1. Introduction

A proposal which has frequently been made in connection with variable frequency synchrotrons is that the wide frequency range required should be generated by a number of sweeping oscillators, each one covering part of the total range. The proposal has never been considered very seriously because of lack of assurance that the essential conditions could be achieved satisfactorily. The difficulty is that, in order to avoid excessive particle loss on switching from one oscillator to the next, the frequency, phase and amplitude discontinuities must be kept small.

In connection with the problem of extracting protons from the Birmingham synchrotron the necessity arose for a similar transition from one oscillator to another ; in this case from an oscillator varying at a rate of 30 Mc sec^{-2} to a fixed oscillator, set to any value in the range 2 to 10 Mc. A frequency jump of 1 part in 10 000 and phase and amplitude jumps of several degrees and several per cent respectively can be tolerated at the transition. These requirements are no less stringent than would apply in the general case referred to above ; moreover, the accuracy of timing for the transition must be very much greater because of the extreme difference in initial and final frequency-time laws (i.e. linear to fixed frequency). The successful results achieved with the simple circuit arrangement to be described here lead to the conviction that the switching of sweeping oscillators in general is quite simple, provided that suitable triggers are available to initiate the actual transitions. Such triggers must be capable of being set to occur when the oscillators are overlapping

* Now at the University of Melbourne.
† Communicated by Professor P. B. Moon, F.R.S.

in frequency within the required tolerance. If the reproducibility of the oscillator laws and the nature of the frequency-control mechanism allow this to be achieved the rest is not difficult. That this is not an excessive requirement of a mechanically tuned system is shown by the fact that, in the equipment under discussion, the short term fluctuations in frequency difference at the transition proved to be of the order of one part in 100 000 of the actual frequency.

It is possible that a beat-frequency discriminator could be employed to determine the transition-switching automatically; however, with a sweep rate of 30 Mc sec^{-2} at a frequency of (say) 2 Mc, the zero-beat pattern is so broad that it would be difficult to devise a waveform discriminator which would give the required time-accuracy.

The derivation of accurate trigger-signals from a rotating condenser (Oliphant 1947 and Hibbard 1950) on the other hand, presents no particular problems and, as in this case the facility was required for other purposes as well, the question of an alternative did not arise. Details of the rotating condenser and trigger pick-up (Caro and Hibbard 1952) will not be described here.

§ 2. General Description

The general method for ensuring accurate phasing of the variable and fixed oscillators at the instant of switching is one which has been used successfully for the instantaneous measurement of the variable-frequency output of the beat oscillator and is described fully in Hibbard and Caro (1952). Briefly, the variable-frequency signal is fed into the 'fixed' oscillator and, as it approaches the 'fixed' frequency, pulls it into synchronism. At the instant of equality of frequency the 'fixed' oscillator is released and carries on at its natural frequency. At the same time the output is switched from one oscillator to the other. Transient switching impulses are of no consequence; however, on the removal of the synchronizing signal, the 'fixed' oscillator decays to a lower level. In order that the final output should remain constant this signal is passed through a limiting buffer, which limits it to a square wave from which the fundamental component is extracted by a tuned circuit. The arrangement of the various units is shown in fig. 1.

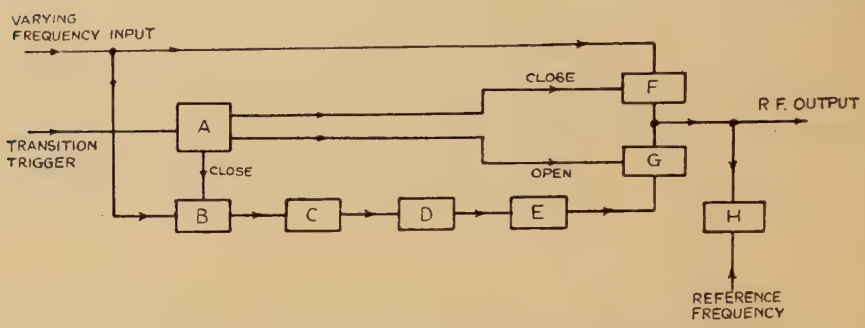
A is a waveform generator which provides the appropriate gating signals to the gates B, F and G. B and F are closed and G is opened at the instant of transition. C, D and E are the synchronizing amplifier, oscillator and limiting buffer respectively. H is a frequency meter which is similar in principle to that described by Hibbard and Caro (1952) and shows any difference in amplitude and frequency at the instant of transition.

§ 3. Circuit Details

The complete circuit is shown in fig. 2. V1 is a bistable circuit which is tripped by the transition trigger and gates the suppressor grids of the gate valves V2, V6 and V7. It is reset by another pulse before the next

cycle is to occur. The gate valves have short suppressor-bases and internal diodes connected to the suppressor grids. Gating waveforms are supplied via resistive potential-dividers, arranged to pull up against these diodes to define the suppressor potential and give a low impedance to earth. The bistable anodes have to control a total of five potential-dividers, each with speed-up condensers, and it is desirable for the gating waveforms to have sharp fronts. The double cathode-follower V10 is necessary to enable the desired speeds to be achieved. The gate circuit

Fig. 1

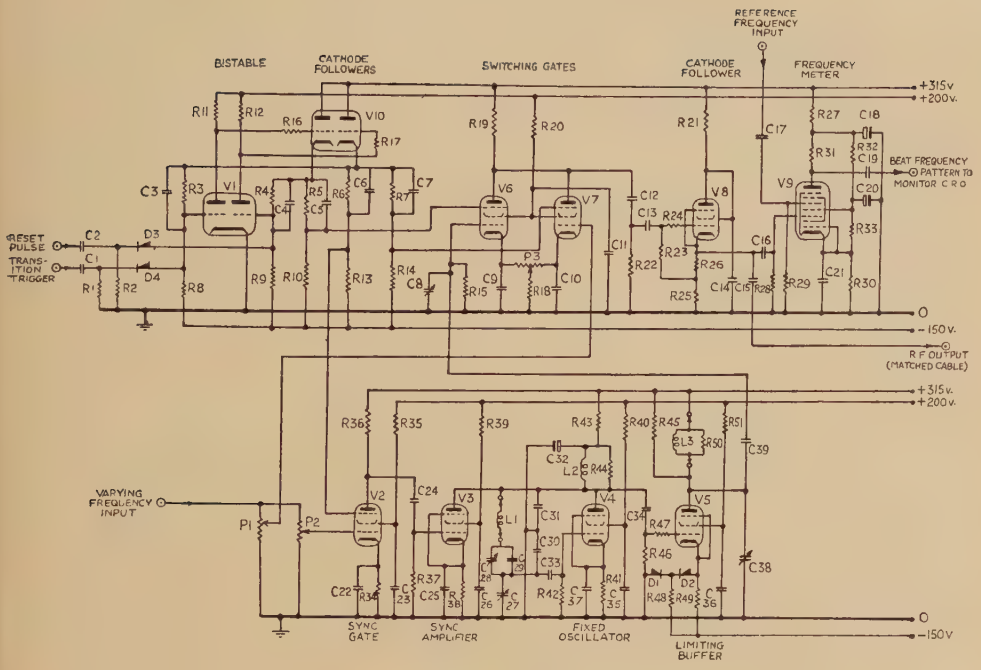


Block diagram of oscillator switching system.

A. Bistable controlling gates.
B. Synchronizing gate.
C. Synchronizing amplifier.
D. Fixed oscillator.
E. Limiting buffer.
F. Varying-frequency gate.
G. Fixed-frequency gate.
H. Frequency meter.

of V2 is very simple and no attempt is made to prevent gating transients from occurring. In the original circuit a twin-tube balanced gate with transformer output was employed in order to eliminate transients, but it was found possible to reduce their effects to negligible proportions by very simple means, so the simpler gate circuit was adopted. The main effect of the gating transient is to throw a large positive pulse on the grid of V3, which in turn depresses the anode potential of V4. Although the type of oscillator shown is relatively insensitive to valve conditions, the frequency stability which is required is such that any disturbance should be kept as small as possible, both as regards magnitude and duration. The magnitude depends mostly on the nature of the pulse on the grid of V3 and the duration depends on the recovery time-constant associated with V4. The first has been reduced by using a small coupling time-constant between V2 and V3, and the second by the choice of shunt-feed components, L_2, R44 in the anode circuit of V4. Any disturbance at the anode decays as a damped sinusoid determined by L_2, C31 and R44. The resultant effect of the switching transient on the anode potential of V4 is thus a disturbance of a few volts persisting for about one μsec. The effect on the oscillator frequency is unobservable.

Fig. 2



Complete circuit.

Resistors

R1, R2	22 k Ω	¼w	R24, R25	1 k Ω	¼w	R39, R40	1 k Ω	¼w
R3 to R10	220 k Ω	¼w	R26	220 Ω	¼w	R41	330 Ω	¼w
R11, R12	50 k Ω	2w	R27	10 k Ω	1w	R42	47 k Ω	¼w
R13, R14	270 k Ω	¼w	R28, R29	47 k Ω	¼w	R43	47 k Ω	2w
R15	10 k Ω	¼w	R30	180 Ω	¼w	R44	10 k Ω	¼w
R16, R17	3·3 k Ω	⅛w	R31	100 k Ω	¼w	R45	47 k Ω	½w
R18	110 Ω	¼w	R32	33 k Ω	3w	R46	22 k Ω	¼w
R19	1·5 k Ω	¼w	R33	1 k Ω	¼w	R47	330 Ω	¼w
R20	12 k Ω	1w	R34, R35	220 Ω	¼w	R48	22 k Ω	2w
R21	10 k Ω	2w	R36	1·8 k Ω	¼w	R49	68 k Ω	½w
R22	1·5 k Ω	¼w	R37	1·5 k Ω	¼w	R50	Adjust for coil Q of 4	
R23	470 k Ω	¼w	R38	100 Ω	¼w	R51	1 k Ω	¼w

Condensers

C1, C2	1000 pF	350 v	C18	8 μF	450 v	C28	3–28 pF	variable
C3, C4	10 pF	350 v	C19	0·1 μF	350 v	C29	50 pF	350 v
C5, C6, C7	20 pF	350 v	C20, C21	25 μF	25 v	C30, C31	250 pF	350 v
C8	1000 pF	variable	C22	0·04 μF	150 v	C32	2 μF	450 v
C9, C10	0·25 μF	150 v	C23	0·01 μF	350 v	C33, C34	120 pF	350 v
C11	0·01 μF	350 v	C24	500 pF	350 v	C35, C36	0·01 μF	350 v
C12, C13	500 pF	350 v	C25	0·1 μF	350 v	C37	0·1 μF	350 v
C14, C15	0·1 μF	350 v	C26	0·01 μF	350 v	C38	3–60 pF	variable
C16, C17	1000 pF	350 v	C27	3 pF	variable	C39	15 pF	350 v

Valves

V1, V10	12AT7
V2, V6, V7	6F33
V3, V4, V5, V8	EF91
V9	EQ80
D1, D2	Two B.T.H. Germanium diodes type CG1-C in parallel, chosen to give low forward resistance.
D3, D4	B.T.H. Germanium diode CG4-C

Miscellaneous

P1, P2	300 ohm 1w potentiometer	L1, L3	Plug in coils to suit each range
P3	25 ohm potentiometer	L2	0·25 mH r.f. choke

The oscillator is of the high-stability Clapp type (Clapp 1948, Roberts 1948) in which the principal frequency-determining elements L_1, C28, C29 are isolated from the valve by the large capacitances C30, C31. The basic circuit as well as the mechanism of synchronization are discussed at some length in Hibbard and Caro (1952). Provided that the rate of variation of frequency is not excessive in relation to the Q of the fixed oscillator, the latter will pull into synchronism shortly before the frequencies become equal, and when they are equal the injected synchronizing current will be very nearly in phase with that supplied by the oscillator valve V4. The phase relations at various points in the oscillator circuit are thus very nearly identical with those which exist in the free running condition, the essential difference being that the level of oscillation is higher. On the removal of the synchronizing current the oscillation decays to the free running level with a time constant of some tens of μsec. No frequency or phase readjustments of significance are observed. The change in oscillation level depends on the ratio of r.f. currents from V3 and V4. This must be sufficiently great to ensure phase-locking before frequency equality is reached and will depend on the sweep rate and the oscillator Q. In the worst cases readjustment of some tens of per cent will occur. With the aid of the limiting buffer V5 the output variation due to this is reduced to a permissible amount.

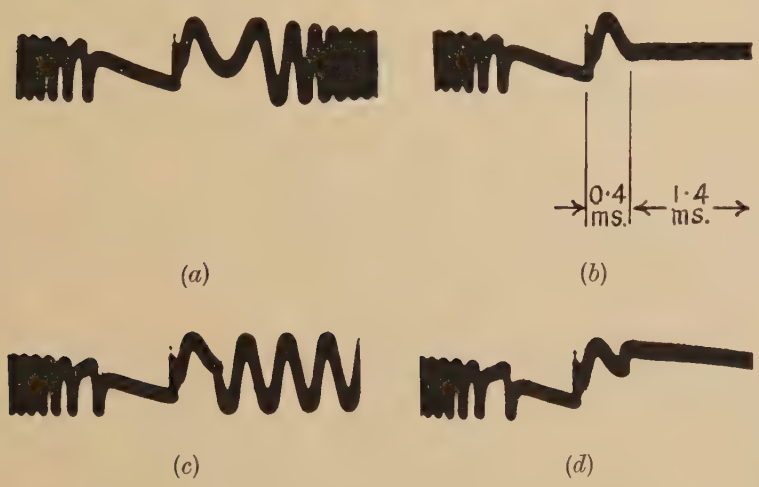
The limiting action of V5 is obtained with the aid of the germanium diodes D1 and D2. If the cathode current is less than 2·2 mA the cathode potential drops below ground and disengages the diode D2. If it is above 9 mA the cathode potential rises above ground and disengages D1. It is only between these two currents that V5 has a sensibly earthed-cathode and functions as an amplifier. With either diode disengaged the valve functions as a cathode follower and has a relatively low grid-anode transconductance. With a sufficiently large input voltage the valve thus passes an essentially square-wave anode current and in neither limiting condition does it load the input circuit. Because of diode impedance and cathode-circuit stray capacitances the limiting action is not perfect but it is adequate for the purpose.

The tuned circuit L_3, C_{38} in the anode of V5 is tuned to the fixed frequency and removes the unwanted harmonic components from the limited output. The tuned circuit is damped in order to lower its Q; a value of 4 gives a suitable compromise between waveform requirements and ease of tuning. If the circuit is mistuned slightly it introduces a phase shift without much change of amplitude and can be used to adjust the phase shifts in the two paths to equality. The transition gates V6, V7 have a common cathode resistor R18 and balancing potentiometer P3. By adjustment of P3 and the correct choice of the speed-up condenser C5, switching transients at the common-anode point can be reduced to a negligible magnitude. These adjustments are best made with the aid of a cathode ray oscilloscope connected to the anodes. The short time-constant circuit R22, C12, coupling to the output cathode-

follower, is inserted in order to minimize the effect of any gating transient which is left uncorrected. It is probable that the switching from one r.f. waveform to the other can be effected without seriously disturbing even one cycle. For the immediate purpose for which the equipment was developed, i.e. for accelerating protons, transient and r.f. disturbances lasting for some μsec could be tolerated and, for this reason, there has been no need to investigate these possibilities more fully.

The adjustment of frequency, phase and time transition in order to achieve smooth hand-over from one signal to the other is done with the aid of V9, a Mullard type EQ80 Nonode. In this valve the output signal

Fig. 3



Frequency monitor waveforms.

(*a*) Linearly varying frequency passing through reference frequency.
(*b*) Beat pattern showing transition to fixed frequency equal to reference frequency.
(*c*) Beat pattern showing transition to fixed frequency differing from reference frequency.
(*d*) Beat pattern showing transition to fixed frequency equal to reference frequency, but with a phase or amplitude discontinuity.

is fed to one grid and a signal from a reference oscillator is fed to another. The anode current is proportional to the product of the two signals, and a beat pattern is obtained which can be viewed on an oscilloscope. This pattern is exactly similar to that described in Hibbard and Caro (1952), and, as described there, a stable repetitive pattern can be achieved by locking the phase of the reference signal to the varying-frequency signal and releasing it just prior to the transition. The types of CRT pattern which can be obtained are shown in fig. 3. In this, fig. 3 (*a*) shows the beat pattern obtained without switching from the varying to the fixed oscillator. The pattern consists of two regions, in the first of which the

reference oscillator can be seen being pulled into synchronism with the varying oscillator and, in the second, it is allowed to run freely. The axis of symmetry of the latter is the instant of equality of the varying and reference frequencies, and is the point at which the switching should be made to occur. In fig. 3 (*b*) the pattern, due to switching at this point from the varying to a fixed frequency identical with the reference frequency, is shown. The effect of an error in the fixed frequency is shown in fig. 3 (*c*). The straight line after switching is now replaced by a sine wave. These photographs were taken at a frequency of 7 mc, and the error demonstrated is about 3 kc, or 1 part in 2500.

The effect of a phase or amplitude discontinuity at switching is shown in fig. 3 (*d*). In practice the requirement is to eliminate the step and achieve the pattern of 3 (*b*), and to have no amplitude discontinuity. If the phase difference through the circuit has been reduced to zero, achieving pattern 3 (*b*) is all that is required. If, however, the phase difference is not known, it is necessary to have an additional facility for observing amplitude equality and adjusting it separately. This can be done easily with the aid of a diode rectifier on the r.f. output, giving a signal on a CRT, which can be adjusted to show no step at the switching instant. When this condition is achieved any step observed in the beat pattern 3 (*d*) will be due to a phase error and can be eliminated by detuning the buffer anode circuit.

Acknowledgments

The work described in this paper has been carried out as part of the programme for building the Birmingham Proton Synchrotron which is financed principally by the Department of Scientific and Industrial Research. One of the authors (D. E. C.) is indebted to the Royal Commission for the exhibition of 1851 for the provision of an overseas research scholarship. The other author (L. U. H.) is indebted to the Walter and Eliza Hall Trust and the Australian National University for the provision of fellowships, during the tenure of which the work was carried out.

References

Caro, D. E., and Hibbard, L. U., 1952, *J. Sci. Instrum.*, **29**, 403.
Clapp, J. K., 1948, *Proc. I.R.E.*, **36**, 356.
Hibbard, L. U., 1950, *Nucleonics*, **7**, No. 4, 30.
Hibbard, L. U., and Caro, D. E., 1952, *J. Sci. Instrum.*, **29**, 366.
Hibbard, L. U., *et al.*, *A servo-system for accurate speed control* (to be published).
Oliphant, M. L., *et al.*, 1947, *Proc. Phys. Soc.*, **59**, 666.
Roberts, R. A., 1948, *Proc. I.R.E.*, **36**, 1261.

CI. *A Variational Formulation of the Multi-Stream Electrodynamic Field Equations*

By P. N. Butcher
Telecommunications Research Establishment, Ministry of Supply, Great Malvern*

[Received April 29, 1953]

Summary

A non-relativistic approximation to Dirac's new variational formulation of the single-stream electrodynamic field equations is presented using 3-vector notation and m.k.s. units throughout. The Hamilton–Jacobi theory of a rotational space charge stream is developed. The variational formulation is generalized to the multi-stream case—both for a finite number and a non-denumerably infinite number of streams.

§ 1. Introduction

In two recent papers Dirac (1951, 1952) has presented a variational formulation of the single-stream electrodynamic field equations using 4-tensor notation. Although this notation is the most elegant in which to present a relativistically invariant theory, it is too condensed to allow the non-specialist easily to appreciate the physical content of the equations. The first object of this paper will be, therefore, to present a non-relativistic approximation to this formulation using the more familiar 3-vector notation and m.k.s. units throughout, and to relate it to already well-known results in theoretical electrodynamics. The variational formulation is accomplished by the introduction of new dynamical variables to describe the space charge. This is facilitated by the use of the Hamilton–Jacobi theory of a space charge stream and the second object of this paper will be, therefore, to present this theory. Finally we shall generalize the formulation to the multi-stream case for both a finite number and a non-denumerably infinite number of streams.

While this paper was in preparation a paper by Buneman (1952) appeared which to some extent accomplishes our first task. However, Buneman is still primarily concerned with the relativistically invariant theory and it is felt that our initial discussion does not lose value thereby. The relativistically invariant theory can also be developed in 3-vector notation but to do so would necessitate an *à priori* choice of field variables (by appropriate juggling with the familiar relativistic factor $\sqrt{(1-\mathbf{v}^2/c^2)}$) which might appear to the non-specialist merely to complicate an otherwise simple discussion. Moreover, in applications to most electronic devices we would be concerned with the non-relativistic approximation.

*Communicated by the Author.

§ 2. Variational Formulation of Maxwell's Equations

In non-relativistic treatments the field variables of a single-stream electrodynamic system are usually taken to be the space charge density ρ and the components of the space charge velocity $\mathbf{v}$, the electric vector $\mathbf{E}$ and the magnetic induction vector $\mathbf{B}$. These variables are related by the non-relativistic electrodynamic field equations

$$\nabla\wedge\mathbf{E}+\dot{\mathbf{B}}=0, \qquad (1)$$

$$\nabla\cdot\mathbf{B}=0,$$

$$\nabla\wedge\mathbf{B}-\frac{1}{c^2}\dot{\mathbf{E}}=\mu\rho\mathbf{v}, \qquad (2)$$

$$\nabla\cdot\mathbf{E}=\rho/\epsilon,$$

$$\frac{D}{Dt}(m\mathbf{v})=-|e|[\mathbf{E}+\mathbf{v}\wedge\mathbf{B}], \qquad (3)$$

in which $D/Dt=\partial/\partial t+(\mathbf{v}\cdot\nabla)$, $|e|$ is the modulus of the electronic charge, m is the electronic rest mass and ϵ and μ are the dielectric constant and magnetic permeability of free space. These equations differ from the relativistically invariant equations only in the appearance of the non-relativistic momentum $m\mathbf{v}$ in the Lorentz equation instead of the relativistic momentum $m\mathbf{v}/\sqrt{(1-\mathbf{v}^2/c^2)}$ i.e. we are disregarding the dependence of the electronic mass on velocity which should of course only lead to errors of a few per cent if the kinetic energy of the electrons is less than about 25 kv.

The general solution of (1) is

$$\mathbf{E}=-(\dot{\mathbf{A}}+\nabla\phi), \qquad (4\,a)$$

$$\mathbf{B}=\nabla\wedge\mathbf{A}, \qquad (4\,b)$$

where $\mathbf{A}$ and ϕ are the so-called vector and scalar potentials of the field. We note that these potentials are not uniquely determined by the electro-magnetic field. Thus, if $\mathbf{E}$ and $\dot{\mathbf{B}}$ can be derived from $\mathbf{A}$ and ϕ through (4), then inspection of these equations shows that they can also be derived from

$$\mathbf{A}^0=\mathbf{A}+\nabla\chi, \qquad (5)$$

$$\phi^0=\phi-\dot{\chi}.$$

Such a transformation of potentials is called a gauge transformation and χ is called the gauge function.

Substituting (4) in (2) we find that they become the Euler equations derived from the Lagrangian density

$$\mathscr{L}=\tfrac{1}{2}\left(\frac{1}{c^2}\mathbf{E}^2-\mathbf{B}^2\right)-\mu\rho(\phi-\mathbf{v}\cdot\mathbf{A}), \qquad (6)$$

when $\mathbf{A}$ and ϕ are varied and ρ and $\mathbf{v}$ are regarded as known (Goldstein 1951). This well known result takes us close to our goal but there is more to be done yet because the Lorentz equation is not included in the above variational principle and we do not know *à priori* the field variables

ρ and $\mathbf{v}$. A solution to the problem is obtained when we eliminate the Lorentz equation from the set of field equations by using it to define the dynamical variables of the space charge (i.e. its velocity and kinetic energy per electronic charge) in terms of the auxiliary potentials and further auxiliary variables.

§ 3. Inclusion of the Equations of Motion in the Variational Principle

An immediate consequence of the Lorentz equation is Lagrange's theorem (Gabor 1945) which states that the circulation of the canonical momentum round any circuit which moves with the space charge is constant, i.e.

$$\frac{D}{Dt}\oint_s \mathbf{p}\,.\,\mathbf{ds} \equiv \frac{D}{Dt}\int_S \mathbf{w}\,.\,\mathbf{dS} = 0, \qquad \ldots\ldots \quad (7)$$

where

$$\mathbf{p} = m\mathbf{v} - |e|\,\mathbf{A},$$

which we shall call the canonical momentum, and

$$\mathbf{w} = \nabla \wedge \mathbf{p} = \nabla \wedge m\mathbf{v} - |e|\,\mathbf{B},$$

which we shall call the vorticity, s is any circuit which moves with space charge and S is any surface bounded by s. We note that $\mathbf{w}$ has a direct physical interpretation.

If $\mathbf{w} \not\equiv 0$ the field is said to be rotational and we may introduce the following useful concepts which are already well established in hydrodynamical theory (Lamb 1924). First of all we define a vortex line as a line drawn in the space charge such that the tangent at any point on it is parallel to $\mathbf{w}$. Thus the vortex lines are the double infinity of lines defined by the differential equations

$$\frac{dx}{w_x} = \frac{dy}{w_y} = \frac{dz}{w_z}.$$

Next we define a vortex tube as a tube formed by a single infinity of vortex-lines all of which pass through a simply connected circuit in the space charge. Finally we define the strength of the vortex tube as the circulation of $\mathbf{p}$ round any simple-connected circuit encircling the tube.

Using (7) it is easy to show that vortex lines and tubes move with the space charge and the strength of a tube is constant in time. These properties allow us to obtain a simple description of the space charge by introducing labels to distinguish the vortex lines. This is most easily done at one instant of time by setting up a co-ordinate mesh on a surface which intersects all the vortex lines once and once only. If ξ and η are the co-ordinates on this surface then we may label each vortex line by the corresponding ξ and η. We fix the labelling for all time by stipulating that each vortex line keeps its initial labelling throughout its subsequent motion. There is a vortex line through each point of the space charge at each instant of time so that we can associate definite

values of ξ and η with each point in the space charge at each instant of time, i.e. ξ and η can be regarded as field variables. Since the vortex lines move with the space charge and keep the same labels for all time it follows that

$$\frac{D\xi}{Dt}=\frac{D\eta}{Dt}=0. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (8)$$

We now turn our attention to the task of expressing the dynamical variables of the space charge in terms of the auxiliary potentials and the vortex line labels. The intersection of the surfaces $\xi=$constant and $\eta=$constant, is a vortex line. Hence $\mathbf{w}$ is perpendicular to the normals to both these surfaces so that

$$\mathbf{w}=\lambda\nabla\xi\wedge\nabla\eta, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

where λ is another scalar variable which depends on the initial co-ordinate system used in the labelling. In the appendix it is shown that by appropriate choice of this system we can make $\lambda=1$. Supposing this to have been done and using the vector identity $\nabla\wedge\nabla\eta=\nabla\wedge(\xi\nabla\eta)$, we have

$$\mathbf{w}\equiv\nabla\wedge\mathbf{p}=\nabla\xi\wedge\nabla\eta\equiv\nabla\wedge(\xi\nabla\eta).$$

Hence

$$\mathbf{p}=\xi\nabla\eta+\nabla S, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (10\,a)$$

where S is a further scalar variable. Now, using the vector identity

$$(\mathbf{v}\,.\,\nabla)\mathbf{v}=\nabla(\tfrac{1}{2}\mathbf{v}^2)-\mathbf{v}\wedge\nabla\wedge\mathbf{v},$$

we can easily write the Lorentz eqn. (3) in the form

$$\dot{\mathbf{p}}-\mathbf{v}\wedge\nabla\wedge\mathbf{p}=-\nabla(T-|e|\phi),$$

where $T=(m\mathbf{v})^2/2m$ is the kinetic energy of the space charge per electronic charge. Substituting (10 a) in this equation, and using (8) we find

$$\nabla(\xi\dot{\eta}+\dot{S})=-\nabla(T-|e|\phi).$$

Integrating, and absorbing the unknown function of t in S (which we see from (10 a) is arbitrary to this extent) we have

$$\xi\dot{\eta}+\dot{S}=-(T-|e|\phi). \quad . \quad . \quad . \quad . \quad . \quad . \quad (10\,b)$$

Rewriting (10) in the form

$$m\mathbf{v}=|e|\mathbf{A}+\xi\nabla\eta+\nabla S, \quad . \quad . \quad . \quad . \quad . \quad . \quad (11)$$
$$T=|e|\phi-\xi\dot{\eta}-\dot{S},$$

we have the required equations defining the dynamical variables of the space charge in terms of the auxiliary potentials, the vortex line labels and the further auxiliary variable S which we shall call the action function. The justification of this terminology will be seen in § 4. These auxiliary field variables will be related by the equation $T=(m\mathbf{v})^2/2m$ which is necessitated by the definition of T. The eqns. (8), (11), and $T=(m\mathbf{v})^2/2m$ provide us with the general solution of the Lorentz eqn. (3). Buneman (1952) points out that (11) is essentially Clebsch's transformation (see Lamb 1924, article 167).

The fundamental field variables of the single stream electrodynamic system may therefore be taken to be $\mathbf{A}$, ϕ, ρ, ξ, η and S in terms of which $\mathbf{E}$ and $\mathbf{B}$ are given by (4) and $m\mathbf{v}$ and T are given by (11). The field equations relating these variables are (2), (8) and $T=(m\mathbf{v})^2/2m$. These field equations are just the Euler equations derived from the Lagrangian density

$$\mathscr{L}=\tfrac{1}{2}\left(\frac{1}{c^2}\mathbf{E}^2-\mathbf{B}^2\right)-\frac{\mu}{|e|}\rho\left\{T-\frac{1}{2m}(m\mathbf{v})^2\right\}, \quad . \quad . \quad (12)$$

when $\mathbf{A}$, ϕ, ρ, ξ and η are varied while variation of S leads to the equation of conservation of charge which is a consequence of (2).

If $\mathbf{w}\equiv 0$ the stream is said to be irrotational. This case, which is obtained by putting both ξ and η equal to zero in the above, was considered by Dirac in his first paper and had previously been considered in 3-vector notation by Cherry (1943).

§ 4. Hamilton–Jacobi Theory of a Rotational Space Charge Stream

In this section we shall use the notation common in theoretical dynamics: denoting the position vector $\mathbf{r}$ of an infinitesimal element of space charge by q_α and its conjugate momentum $\mathbf{p}$ by p_α.

The canonical co-ordinates (q_α, p_α) of each infinitesimal element of space charge are given as functions of the time and 6 canonical constants of the motion of that element (Q_α, P_α) by the canonical transformation

$$p_\alpha=\frac{\partial}{\partial q_\alpha}S(q_\alpha, P_\alpha, t), \quad . \quad . \quad . \quad . \quad . \quad . \quad (13\,a)$$

$$Q_\alpha=\frac{\partial}{\partial P_\alpha}S(q_\alpha, P_\alpha, t), \quad . \quad . \quad . \quad . \quad . \quad (13\,b)$$

where S is a complete integral of the Hamilton–Jacobi equation

$$H\left(q_\alpha, \frac{\partial S}{\partial q_\alpha}\right)+\frac{\partial S}{\partial t}=0, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (14)$$

with the additive constant set equal to zero (see Goldstein 1951, chapter 9). Q_α and P_α depend on the element of space charge considered, that is on q_α and t so that

$$Q_\alpha=Q_\alpha(q_\alpha, t) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (15\,a)$$

and

$$P_\alpha=P_\alpha(q_\alpha, t). \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (15\,b)$$

By solving (15 *b*) for q_α and substituting in (15 *a*) we can express Q_α as a function of P_α and t, but t cannot appear explicitly because Q_α and P_α are constants of the motion. Hence Q_α is a function of P_α only which we can always express in the form

$$Q_\alpha=\frac{\partial S^0}{\partial P_\alpha}-\xi\frac{\partial \eta}{\partial P_\alpha}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (16)$$

where S^0, ξ and η are functions of P_α (cf. the derivation of eqn. (10 *a*)

or see Lamb 1924, article 167). Substituting (15 *b*) into S and taking total derivatives we obtain

$$\frac{dS}{dq_\alpha}=\frac{\partial S}{\partial q_\alpha}+\Sigma_\beta Q_\beta\frac{\partial P_\beta}{\partial q_\alpha},$$
$$\frac{dS}{dt}=\frac{\partial S}{\partial t}+\Sigma_\beta Q_\beta\frac{\partial P_\beta}{\partial t}.$$

Hence, using (16), we have

$$\left.\begin{aligned} p_\alpha=\frac{\partial S}{\partial q_\alpha}&=\frac{d}{dq_\alpha}(S-S^0)+\xi\frac{d\eta}{dq_\alpha},\\ \frac{\partial S}{\partial t}&=\frac{d}{dt}(S-S^0)+\xi\frac{d\eta}{dt}. \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad (17\,a)$$

Substituting in the Hamilton-Jacobi equation (14) we obtain

$$H\left(q_\alpha,\frac{d}{dq_\alpha}(S-S^0)+\xi\frac{d\eta}{dq_\alpha}\right)+\frac{d}{dt}(S-S^0)+\xi\frac{d\eta}{dt}=0 \quad . \quad . \quad (17\,b)$$

Finally, since ξ and η are functions of the constants of the motion P_α only, we have

$$\frac{D\xi}{Dt}=\frac{D\eta}{Dt}=0. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (18)$$

We recognize (18) as the dynamical eqns. (8) and, absorbing the ' additive constant ' S^0 in S and remembering that $H=T-|e|\phi$, we see that (17) comprises the definitions (11) of the dynamical variables of the space charge $m\mathbf{v}$ and T in terms of $\mathbf{A}$, ϕ, ξ, η and S. We note that, although in these equations S is regarded as a function of q_α and t only, it may still be called the action function with some justification. The development of the field equations from this point is trivial. The Lorentz invariant equations will be obtained if the relativistic Hamiltonian is used, but we shall not concern ourselves with writing down the equations for this case.

§ 5. The Multi-Stream Electrodynamic System

We consider first of all the case of a finite number (N) of streams. Let the nth stream have space charge density ρ_n, velocity $\mathbf{v}_n$ and kinetic energy per unit electronic charge T_n. The discussion of § 3 (or § 4) can be carried through for each stream separately and leads to the introduction of an appropriate pair of vortex line labels ξ_n and η_n and an action function S_n to describe the motion of the nth stream. The fundamental field variables of the system may therefore be taken to be $\mathbf{A}$, ϕ, ρ_n, ξ_n, η_n and S_n in terms of which $\mathbf{E}$ and $\mathbf{B}$ are given by (4) and $\mathbf{v}_n$ and T_n are given by (11) with suffix n introduced. The field equations relating these variables are: (2) with $\rho\mathbf{v}$ and ρ replaced by $\Sigma_n\rho_n\mathbf{v}_n$ and $\Sigma_n\rho_n$, (8) with ξ and η replaced by ξ_n and η_n, the equations of conservation of charge:

$$\dot{\rho}_n+\nabla\cdot(\rho_n\mathbf{v}_n)=0, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (19)$$

and the equations $T_n=(m\mathbf{v}_n)^2/2m$. These field equations are just the

Euler equations derived from the Lagrangian density

$$\mathscr{L}=\tfrac{1}{2}\left(\frac{1}{c^2}\,\mathbf{E}^2-\mathbf{B}^2\right)-\frac{\mu}{|e|}\Sigma_n\rho_n\left\{T_n-\frac{1}{2m}\,(m\mathbf{v}_n)^2\right\}, \quad . \quad (20)$$

when $\mathbf{A}$, ϕ, ρ_n, ξ_n, η_n and S_n are varied. The irrotational case is obtained by setting both ξ_n and η_n equal to zero for all n.

We consider now the case of a non-denumerably infinite number of streams. In this picture of an electrodynamic system, instead of smearing out the electrons into a finite number of physical ($\mathbf{r}$) space charge streams, we smear them out into a phase ($\mathbf{r}$, $\mathbf{p}$) space charge stream moving with the canonical velocity

$$\left.\begin{aligned}\frac{dq_\alpha}{dt}&=\frac{\partial H}{\partial p_\alpha},\\ \frac{dp_\alpha}{dt}&=-\frac{\partial H}{\partial q_\alpha}.\end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (21)$$

We can break down this sextuple infinity of trajectories into a triple infinity of streams in $\mathbf{r}$ space each moving in accordance with (21) and in each of which $\mathbf{p}$ is a function of $\mathbf{r}$ and t. To do so we prescribe $\mathbf{p}$ at an instant t_0 as a suitable function of $\mathbf{r}$ and a stream labelling 3-vector $\mathbf{n}$ (which takes on a continuum of values) and subsequently allow the points in phase space thus defined to move in accordance with (21). Then we can introduce the charge density in ($\mathbf{r}$, $\mathbf{n}$) space—which we shall denote by $\rho(\mathbf{r}, \mathbf{n}, t)$—to describe the distribution of the phase space charge. We shall denote the velocity and corresponding electronic kinetic energy of the $\mathbf{n}$th stream by $\mathbf{v}(\mathbf{r}, \mathbf{n}, t)$ and $T(\mathbf{r}, \mathbf{n}, t)\equiv(m\mathbf{v})^2/2m$ respectively.

The discussion of § 3 (or § 4) can be carried through for each stream separately and leads to the introduction of an appropriate pair of vortex line labels $\xi(\mathbf{r}, \mathbf{n}, t)$ and $\eta(\mathbf{r}, \mathbf{n}, t)$ and an action function $S(\mathbf{r}, \mathbf{n}, t)$ to describe the motion of the $\mathbf{n}$th stream. The fundamental field variables of the system may therefore be taken to be $\mathbf{A}$, ϕ, ρ, ξ, η and S in terms of which $\mathbf{E}$ and $\mathbf{B}$ are given by (4) and $\mathbf{v}$ and T are given by (11) in which the stream label $\mathbf{n}$ is now implicit. The field equations relating these variables are : (2) with $\rho\mathbf{v}$ and ρ replaced by $\int\rho\mathbf{v}\,d\mathbf{n}$ and $\int\rho\,d\mathbf{n}$, (8) in which the stream label $\mathbf{n}$ is now implicit, the equations of conservation of charge

$$\dot{\rho}(\mathbf{r}, \mathbf{n}, t)+\nabla\cdot[\rho(\mathbf{r}, \mathbf{n}, t)\mathbf{v}(\mathbf{r}, \mathbf{n}, t)]=0; \quad . \quad . \quad . \quad . \quad (22)$$

and the equations $T=(m\mathbf{v})^2/2m$. These field equations are just the Euler equations derived from the variational principle

$$\delta\int d\mathbf{r}\,dt\left[\frac{1}{2}\left(\frac{1}{c^2}\,\mathbf{E}^2-\mathbf{B}^2\right)-\frac{\mu}{|e|}\int d\mathbf{n}\rho\left\{T-\frac{1}{2m}\,(m\mathbf{v})^2\right\}\right]=0, \quad (23)$$

when $\mathbf{A}$, ϕ, ρ, ξ, η and S are varied.

The specification of the $\mathbf{r}$ space streams by defining $\mathbf{p}$ as a function of $\mathbf{r}$ and $\mathbf{n}$ when $t=t_0$ is at our disposal. If we take, for example,

$$\mathbf{p}=\nabla S(\mathbf{r}, \mathbf{n}, t_0),$$

where S is the complete integral of the Hamilton–Jacobi eqn. (14) with

the additive constant set equal to zero, then the canonical momentum in the nth stream will be given as a function of $\mathbf{r}$ and t by

$$\mathbf{p} = \nabla S(\mathbf{r}, \mathbf{n}, t).$$

Thus all the streams are irrotational and we can set both $\xi(\mathbf{r}, \mathbf{n}, t)$ and $\eta(\mathbf{r}, \mathbf{n}, t)$ equal to zero in the above discussion. We see therefore that in the present case we can regard the $\mathbf{r}$ space streams as either rotational or irrotational. The two points of view are equivalent and which we adopt will depend on the particular topic under discussion.

§ 6. Conclusion

It will be observed that the action function plays a prominent part in our analysis. In the single stream case it can be eliminated from the equations by a gauge transformation of the auxiliary potentials with the gauge function $|e|^{-1}S$. This is the course followed by Dirac. However, to do so involves placing a gauge restriction on the potentials which would be inconvenient in many practical applications in which the magnetic field due to the moving space charge is negligible so that the magnetic field is just the known applied field from which $\mathbf{A}$ can be found provided there are no gauge restrictions. Moreover, this elimination is peculiar to the single stream case ; it is not possible to eliminate all the action functions of the multi-stream case.

Direct integration of the electrodynamic field equations is only feasible with very simple boundary conditions. It would be reasonable to attempt instead the solution of a corresponding variational problem since such attempts have had considerable success in other branches of mathematical physics. However, the boundary conditions to electrodynamic problems do not usually give the boundary values of all the field variables. As a consequence the corresponding variational problems will be unusual and the technique of solution may be expected to differ from that traditionally used. There has been little investigation of this topic which appears to hold some hope of solution of the complicated electrodynamical problems relating to electronic devices.

Acknowledgments

Acknowledgment is made to the Chief Scientist, Ministry of Supply, for permission to publish this paper. Crown Copyright reserved. Reproduced by permission of the Controller, H.M. Stationery Office.

References

Buneman, O., 1952, *Proc. Roy. Soc.* A, **215,** 346.
Cherry, T. M., 1943, *Rep. Coun. Sci. Industr. Res. Aust.*, No. MUM–1.
Dirac, P. A. M., 1951, *Proc. Roy. Soc.* A, **209,** 291 ; 1952, *Ibid.*, **212,** 330.
Gabor, D., 1945, *Proc. Inst. Radio Engrs., N.Y.*, **33,** 792.
Goldstein, H., 1951, *Classical Mechanics* (Addison-Wesley Press. Inc.).
Lamb, H., 1924, *Hydrodynamics*, 5th ed. (Cambridge : University Press), Chap. 7.

APPENDIX

For any initial co-ordinate system we have, from (9), for the strength of a vortex tube

$$\Omega = \int_S \lambda \nabla\xi \wedge \nabla\eta \,.\, \mathbf{dS},$$

where S is any cross-section of the tube. If the elementary parallelogram formed on S by the mesh of intersections of the surfaces ξ=constant and η=constant has sides $\mathbf{dl}_1$ and $\mathbf{dl}_2$ along the ξ lines and η lines respectively so that $\mathbf{dS}=\mathbf{dl}_1 \wedge \mathbf{dl}_2$, then we have

$$\begin{aligned}\Omega &= \int_S \lambda(\nabla\xi \wedge \nabla\eta) \,.\, (\mathbf{dl}_1 \wedge \mathbf{dl}_2) \\ &= \int_S \lambda[(\nabla\xi \,.\, \mathbf{dl}_1)(\nabla\eta \,.\, \mathbf{dl}_2) - (\nabla\xi \,.\, \mathbf{dl}_2)(\nabla\eta \,.\, \mathbf{dl}_1)] \\ &= \int \lambda \, d\xi \, d\eta,\end{aligned}$$

where the latter integral is taken over the region of the (ξ, η) plane whose boundary points are the labels of the vortex lines which constitute the vortex tube. This region is time independent because the vortex line labels are. Since Ω does not depend on either the cross section S or the time t chosen for its evaluation it follows that λ must be a function of ξ and η only and not of $\mathbf{r}$ and t, i.e. it is solely determined by the initial labelling. It is now a trivial matter to choose this labelling so that $\lambda=1$. Thus define a new labelling (ξ^0, η^0) by

$$\xi^0 = \xi^0(\xi, \eta),$$

$$\eta^0 = \eta^0(\xi, \eta),$$

where

$$\frac{\partial(\xi^0, \eta^0)}{\partial(\xi, \eta)} = \lambda(\xi, \eta),$$

then, from (9), we have immediately

$$\nabla\xi^0 \wedge \nabla\eta^0 = \mathbf{w}.$$

CII. *The Variation with Distance in the Range* 0–100 km *of Atmospheric Wave-Forms*

By R. B. Morrison
Sir John Cass College, London, E.C.3*

[Received June 12, 1953]

Abstract

The magnitude and wave-form of atmospherics within 100 km of the source have been calculated on the basis of existing information on the lightning discharge. Particular attention is paid to the variation with distance of the wave-form associated with the return stroke. This variation shows several characteristics of interest.

A direct comparison is made between the theoretical results and the experimental observations made on near atmospheric wave-forms.

§ 1. Introduction

Research into the nature of the lightning discharge has been conducted along two definite lines. First, Schonland and others have photographed the lightning flash with a Boys camera and examined the photographic records. The various stages of the discharge process can be studied from the degree of luminosity and time factors involved on the records. Secondly, Appleton and Chapman and others have recorded and examined the electric field changes associated with the lightning discharge. A study of the types, magnitudes and time factors of the field changes can provide further information on the nature of the discharge process. This second approach to the problem is followed here.

The general form of the electric field change associated with near lightning discharges has been known for some time. Appleton and Chapman (1937) pointed out that the near wave-form of the field change consisted of three distinct parts, which they called the ' a ', ' b ' and ' c ' portions. The ' a ' portion is an initial slow increase in the strength of the electric field. This is followed by the ' b ' portion which consists of a rapid and large increase in the electric field. Finally, there is often a subsequent slow increase in the field which constitutes the ' c ' portion. These three portions of the near field change have been identified by Schonland, Hodges and Collens (1938) as corresponding to the leader, return and subsequent strokes of the lightning discharge respectively.

Increased delineation of the ' b ' portion wave-form has enabled a more detailed analysis to be made of the way in which this wave-form changes with distance. It should be possible to correlate the observed change in

* Communicated by R. H. Humphry.

the ' b ' portion wave-form with existing known data on the nature of the lightning return stroke process. Lutkin (1939) has suggested how this may be done from an analysis of the intensity variation on a Boys camera photograph of the return stroke.

The return stroke consists of a streamer of positive electricity moving with high velocity from the earth to the cloud. Schonland, Malan and Collens (1935) have pointed out that this streamer moves with decreasing velocity as it ascends. Bruce and Golde (1941) suggested an equation for this velocity which fits the experimental results fairly well. It is given in eqn. (2) below. Integration of the equation gives 2·7 km for the total height of the discharge channel. Although the cloud base of the cumulonimbus is of the order 1·5 km in temperate latitudes, there are good reasons for believing the discharge centre to be situated some distance within the cloud (Kuettner 1950).

On the basis of experimental observations, Bruce and Golde (1941) and Norinder (1951) have suggested how the return stroke current varies with time. The time scales of their current wave-forms do not agree. This may be due to the fact that the former considered only weak discharges of less than 4000 amp, which may show the shorter wave-form. The formula for the current wave-form given by Bruce and Golde can be made to fit the current wave-form given by Norinder by suitable adjustments in the values of the parameters, as shown below. There is strong evidence, however, for believing that the current wave-form suggested by Norinder fits the majority of lightning discharges.

It is clear that any calculations based on the return stroke process must take into account the changing height of the discharge channel and the appropriate current variation. Previous calculations have often assumed a constant height of too small a value. A satisfactory explanation of the observed change with distance in the ' b ' portion wave-form can be given on the basic properties of the return stroke process outlined above.

§ 2. Experimental Results

A description of the apparatus used for recording the electric field wave-forms of near lightning discharges, together with some typical results, was given by Morrison (1949). Similar results have been reported by Norinder (1951).

Figures 1 to 8 (Plates 40 and 41) show typical field change records arranged for increasing distance from the source within the range 10–100 km. They were obtained during the passage of a violent thunderstorm across London during the late afternoon of June 14, 1948. Although the records were taken in sequence as the storm receded, no direct measurement of the distance was possible. An estimate of the distance was made in most cases by timing the lightning–thunder interval.

Only the ' b ' portion of the electric field change, and the change in the basic wave-form with distance are of interest here. All the records show

a sharp maximum in the electric field 20 μsec after the beginning of the return stroke, and a minimum after 50–100 μsec, depending on the distance.

As the distance of the lightning flash increased, the ratio of the maximum field change to the final net field change increased. The magnitude of the minimum field change decreased as the storm receded and finally appeared as a negative loop (fig. 7). The time of the minimum occurred later as the distance increased. These characteristics of the basic ' b ' portion wave-form need to be correlated with existing information on the return stroke process.

§ 3. Theoretical Considerations

The vertical electric field, E, at a point distant r from a discharge channel where there occurs a change in the electric moment, M, is given by the equation

$$E=(1/r^3)M+(1/cr^2)\,dM/dt+(1/c^2r)\,d^2M/dt^2. \quad . \; . \; . \quad (1)$$

This equation is valid only if the distance r is greater than the height of the discharge channel. It is also assumed that the magnitude of the current along the channel is uniform at any instant.

According to Bruce and Golde (1941), the velocity of the return stroke at any instant may be represented by the equation

$$v=v_0 \exp(-\gamma t), \quad . \; . \; . \; . \; . \; . \; . \; . \quad (2)$$

where $v_0=8\times10^4$ km sec^{-1}, $\gamma=3\times10^4$ sec^{-1}.

Bruce and Golde (1941) have shown that the rate of change of electric moment may be written as

$$dM/dt=2i\int_0^t v\,dt, \quad . \; . \; . \; . \; . \; . \; . \; . \quad (3)$$

where i=value of current in return stroke at time t.

From eqns. (2) and (3), we have

$$dM/dt=2iv_0\{1-\exp(-\gamma t)\}/\gamma. \quad . \; . \; . \; . \; . \quad (4)$$

If the value of the current, i, is known, the value of dM/dt at any instant may be calculated from eqn. (4). From an analysis of surges on transmission lines caused by direct lightning strokes, Bruce and Golde (1941) suggested that the return stroke current may be represented by an equation of the type

$$i=i_0\{\exp(-\alpha t)-\exp(-\beta t)\}, \quad . \; . \; . \; . \; . \quad (5)$$

where $\alpha=4{\cdot}4\times10^4$ sec^{-1}, $\beta=4{\cdot}6\times10^5$ sec^{-1}.

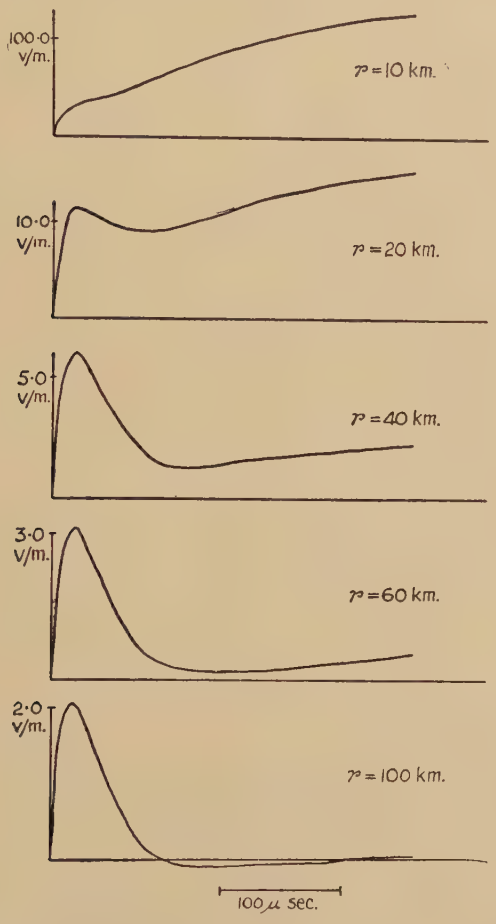
Norinder (1951) suggested that the current takes longer to reach its peak value, and subsequent half value, than given by eqn. (5) with the values of α and β quoted. If, however, values of $\alpha=7\times10^3$ sec^{-1} and $\beta=4\times10^4$ sec^{-1} are substituted in eqn. (5), it is possible to fit Norinder's experimental curve over the range 0–300 μsec fairly well.

Substituting in eqn. (4) for i from eqn. (5) with the new parameters

$$dM/dt=2i_0v_0\{\exp(-\alpha t)-\exp(-\beta t)\}\{1-\exp(-\gamma t)\}/\gamma. \quad . \quad (6)$$

This equation takes into account both the variation of current and the increasing height of the discharge channel with time. From it may be calculated the corresponding values of M and d^2M/dt^2. Substitution in eqn. (1) then gives the value of the electric field. In this way the wave-form of the ' b ' portion may be calculated for different distances from the source, as shown in fig. 9.

Fig. 9



It must be emphasized at this point, that only the basic current and ' b ' portion wave-forms are considered here. In practice no two lightning discharges are alike and there are pulsations superimposed on the basic variation considered. Malan and Collens (1937) attributed such pulsations to the branching of individual strokes. Lutkin (1939) and Norinder (1951) derived wave-forms from a single recorded flash which shows such pulsations. The derived wave-forms thus show certain characteristics not present in other cases, but due to the particular pulsations present in the flash recorded.

It is further assumed that the propagation of the electric field occurs as in free space, and is not subject to modification arising from possible propagation effects. Such assumptions are justified at points close to the lightning flash. At large distances the high frequency components of atmospherics become attenuated, and a wave-form results which cannot be attributed to changes in the discharge channel alone. The function, d^2M/dt^2, which gives the wave-form at large distances must therefore be modified to account for these propagation effects. This is outside the scope of the present paper.

§ 4. Discussion of Results

In fig. 9, the calculated wave-forms for the electric field have been plotted for stated values of r in the range 10–100 km and for $i_0=30$ kilo-amp. For values of r greater than 10 km, a maximum in the value of E occurs at 20 μsec, while a minimum occurs at between 70 and 140 μsec depending on the distance. For values of r greater than 70 km, a negative loop appears in the wave-form. These characteristics agree fairly well with those of the experimental wave-forms in figs. 1 to 8.

As the distance r increases, the ratio of the maximum field, E_R, to the net electrostatic change, E_S, also increases. This is due to the gaining predominance of the radiation term $(1/c^2r)\, d^2M/dt^2$. At 26 km the values of E_R and E_S are equal. This ratio, E_R/E_S, and the distance of 26 km for equality, are independent of the peak current value in the discharge channel. It is thus possible to determine the value of r for a given wave-form by measuring E_R/E_S. Figure 10 shows the variation of E_R/E_S with r. The value of this ratio has been determined where possible from the experimental records and the corresponding value of r found by reference to fig. 10. Such values agree fairly well with the rough estimate from the lightning–thunder interval.

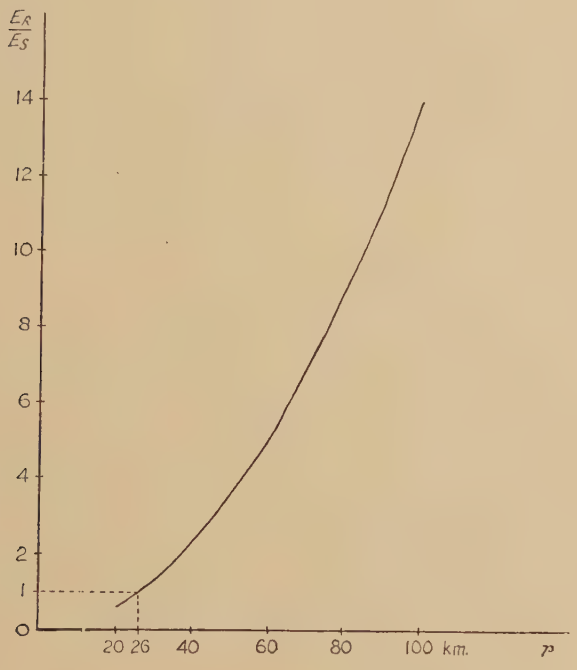
From the computed values of r, it was possible to choose a particular value of i_0 which would give the appropriate measured field values at the given distance. The calculated wave-form could thus be made to fit the experimental wave-form for magnitude as well as form. The values of i_0 required to fit the experimental curves were not unreasonable. The average peak current value is about 20 kilo-amp, which corresponds to a value of i_0 of about 30 kilo-amp. Values of i_0 from 12–80 kilo-amp were required to fit the experimental curves for magnitude.

The value of E_R/E_S determines both the wave-form and the distance of the field change, while i_0 determines the actual magnitudes of E_R and E_S. In figs. 7 and 8, which represent the most distant field changes, E_S can no longer be measured. In this case only an approximate agreement can be found for $r=100$ km and $i_0=80$ kilo-amp. The ratio of the negative loop value to E_R for this distance, however, does not agree with experiment. This may be due to the influence of propagation effects already mentioned.

The calculated wave-forms, together with the values of r and i_0, are given in figs. 1 (a) to 8 (a) alongside the respective experimental wave-forms. The calculated values of E_R and E_S in volts/metre for $i_0=30$ kilo-amp are given in the table. These magnitudes are directly proportional to i_0.

r (km)	10	20	40	60	80	100
E_R(v/m)	—	11·7	5·6	3·4	2·7	2·1
E_S(v/m)	150	18·7	2·4	0·7	0·3	0·15

Fig. 10



If values of $\alpha=4{\cdot}4\times10^4$ sec^{-1} and $\beta=4{\cdot}6\times10^5$ sec^{-1} are taken in eqn. (5), a similar kind of wave-form variation is obtained. The magnitudes of the component features, however, do not fit the experimental data so well. Thus $E_R=E_S$ at 5·5 km instead of at 26 km. Further, the negative loop appears for $r>20$ km instead of $r>70$ km. The maximum field change occurs at 3 μsec instead of 20 μsec, and the minimum at 35 μsec instead of about 100 μsec. While such wave-forms have been observed by Norinder (1934), they form the minority of discharges.

Acknowledgments

In conclusion I should like to record my thanks and appreciation to Dr. F. W. Chapman of King's College, London, for several helpful suggestions and encouragement during the progress of the work. I should also like to thank the Delegacy of King's College, London, for a

Tutorial Studentship from 1946 to 1949 while the experimental work was in progress.

References

Appleton, E. V., and Chapman, F. W., 1937. *Proc. Roy. Soc.* A, **158,** 1.
Bruce, C. E. R., and Golde, R. H., 1941, *J. Inst. E. E.*, **88,** 487.
Kuettner, J., 1950, *J. Met.*, **7,** 322.
Lutkin, F. E., 1939, *Proc. Roy. Soc.* A, **171,** 285.
Malan, D. J., and Collens, H., 1937, *Proc. Roy. Soc.* A, **162,** 175.
Morrison, R. B., 1949, *Ph.D. Thesis*, University of London.
Norinder, H., 1934, *J. Franklin Inst.*, **218,** 717 ; 1951, *International Council of Scientific Unions. Joint Commission on Radio-Meteorology* (1951), p. 17.
Schonland, B. F. J., Hodges, D. B., and Collens, H., 1938, *Proc. Roy. Soc.* A, **166,** 56.
Schonland, B. F. J., Malan, D. J., and Collens, H., 1935, *Proc. Roy. Soc.* A, **152,** 595.

CIII. *A Precision Single-Channel Kicksorter for Coincidence Work*

By Ernst Breitenberger
Cavendish Laboratory, Cambridge*

[Received May 21, 1953]

Summary

A single-channel pulse-height analyser of high speed and precision is described suitable in particular for fast coincidence experiments (resolving time 0·1 μsec). The channel width can be chosen at several values between 2 and 15 v and remains constant to better than 2 parts in 1000 over 10 hours, permitting statistically significant single counts of 10^5. The channel centre is continuously variable within a 50 v range ; its r.m.s. fluctuation during 10 hours is smaller than 70 mv.

§ 1. Introduction

Given electrical pulses of similar shape but different sizes, with some size distribution (their ' spectrum '), a single-channel kicksorter serves to select those pulses whose peak height lies between certain specified limits (the ' channel boundaries ') and indicates the occurrence of such a pulse by a standard signal (the ' yes ').

The stability of the channel boundaries often determines how useful a kicksorter is. One can distinguish two components in boundary migrations (Hutchinson and Scarrott 1951) : changes of channel width (' relative boundary drift ') and of position (' common drift ') ; their effects are quite different.

Relative drift seriously limits the statistical significance of the number N of ' yes ' received through the channel during a given counting time T. Suppose we have a flat (' white ') spectrum and an average channel width w. Then, if during T the width fluctuates *many* times about w, both expectation value and variance of N will remain the same as for a channel of constant width w. But if it changes so slowly that during any time interval of length T it may be regarded as constant, there will be an error in a single count N of approximately N/w times the difference between actual width and w. Thus, whilst the expectation value of N remains of course the same as for a channel of invariable width, the variance of N is now greater by an amount N^2/w^2 times the variance in width. Clearly, unless this increase is negligible, one single count has not the same statistical significance as if it was taken through a stable channel, and several counts of duration T taken at different times cannot be compared

* Communicated by Professor O. R. Frisch, F.R.S.

at all. In particular, for pulses arriving at random, a count N will not be as significant as its magnitude indicates unless the r.m.s. relative drift is smaller than $w/\sqrt{N}$. These arguments hold only for a white spectrum. However, it is easy to see that the conclusions essentially obtain if the channel encloses an arbitrarily shaped portion of the spectrum, provided one disregards extreme cases (e.g. a spectrum consisting of one rectangular line completely inside the channel, or just outside it).

Common drift is another matter. In a white spectrum it cancels out altogether ; if the two boundaries intersect the spectrum at equal heights it cancels out to at least a first approximation, and in all other cases it leads to a smaller error in a single count than that which would be caused by a relative drift twice as large as the common drift. Under these circumstances, it will be possible in almost any experiment to eliminate common drift—although it may be quite large—by appropriately placing the channel.

Thus, long-term stability of channel width must be a main goal in single-channel kicksorter design (cf. Hutchinson and Scarrott 1951, Wilkinson 1949). In most published models (Freundlich, Hincks and Ozeroff 1947, Roulston 1950, Francis, Bell and Gundlach 1951, Eppstein 1951, Van Rennes 1952, Farley 1953, Johnstone 1953) the sorting operation is carried out by two discriminators in an anti coincidence arrangement (see Van Rennes 1952 for references to other systems). Schmidt-type discriminators are used, without questioning at great length the stability of their triggering levels.

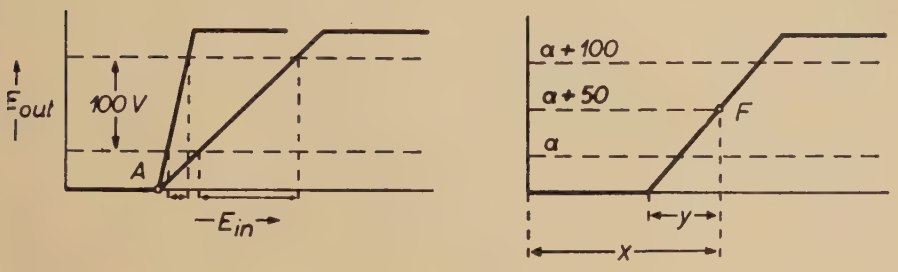
A multivibrator of any kind triggers when the anode currents of the cathode-coupled pair have reached certain values determined by the circuit parameters. Now, at a given current, the grid-to-cathode potential of a valve can vary by as much as 100 mv over periods of the order 10 hours (Gray 1948), the sort of time over which a nuclear experiment is likely to extend. Most of this variation is due only to the cathode (heater voltage fluctuation, ageing, zero-frequency component of flicker effect) and should therefore cancel out between the two grids of either a double triode with common cathode or of any cathode-coupled pair provided the two valves are of the same type, age and place and day of manufacture (Valley and Wallmann 1948, Scarrott 1950). However, some asymmetry —due mainly to varying contact potentials—will remain even in such a balanced pair (Gray 1948). A safe figure for the maximum drift in the difference between the two grid potentials during 10 hours appears to be 50 mv. The triggering point of a multivibrator will be uncertain by this amount ; in the mentioned kicksorters, where the channel is defined more or less *directly* by the discriminator levels, the result is a corresponding uncertainty in both channel width and position.

§ 2. Principle of Operation

In the instrument described here, relative drift is minimized by putting the two discriminator levels so far apart (100 v) that the 50 mv-uncertainty becomes very small compared with the gap between them.

The necessary signal expansion is an easy task because only a narrow slice of each pulse is of interest to the sorter : that overlapping the channel. This slice is selected by a non-overloading biased amplifier with a characteristic as in fig. 1 (*a*) and amplified g times so that it fits in the gap between the two discriminators (similar Van Rennes 1952, Farley 1953,

Fig. 1

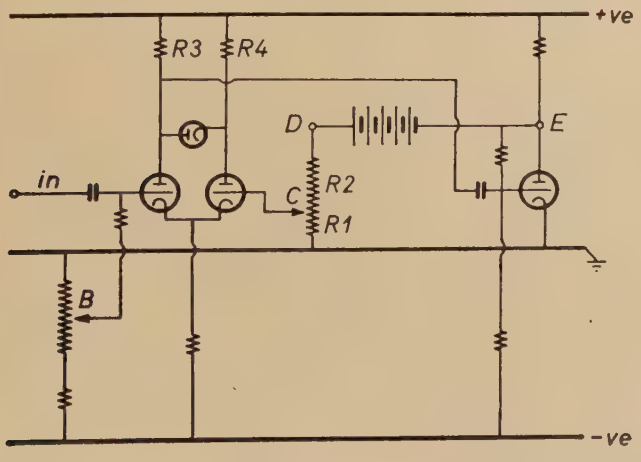


(*a*) (*b*)

Johnstone 1953 ; ' window-amplifier '). Then $w=100/g$, and if we want w to be adjustable from 2–15 v (assuming that the input pulses lie in the amplitude range 0–50 v, as usual with nuclear amplifiers) the required values of g lie between 50 and 6·67.

A suitable biased amplifier is sketched in fig. 2. It accepts positive input pulses ; negative feedback from E to C ensures linearity and gain stability ; overloading will not be harmful because the input cathode follows whereas the last valve is turned off. Either D or E may serve as output terminal.

Fig. 2



Fundamental circuit of biased amplifier.
$R_4=R_3$.

The purpose of the diode is to unhook the first anode, and set the feedback loop in operation, at precisely the moment when a pulse has raised the input grid to the potential V_C. Thereafter, the negative feedback makes the grid at C move very close to the input grid, hence the waveform to be amplified appears at C. Since the output pulse appears at D, the gain is $g=(R_1+R_2)/R_1$.

The action of the diode also defines the bias as V_C-V_B. This difference depends on the chosen gain, because always $V_C=V_D/g$. Thus the situation in fig. 1 (*a*), where the bias is the same for different gains, would require $V_D=0$. But there the channel position depends on the chosen width, which is somewhat inconvenient ; we would rather have a situation as in fig. 1 (*b*), where a change of gain turns the characteristic about the channel centre F. Then the bias must be $x-y=x-(a+50)/g$. Recalling that it is given by $-V_B+V_D/g$, one sees that the choice $V_D=-(a+50)$ leads to the desired arrangement and gives at the same time $x=-V_B$ so that the channel centre position can simply be read off the potentiometer at B.

However, any drift of V_E will appear at D as well and there violate the condition $V_D=-(a+50)$, if a is a *constant*. Long-term drifts are prevented effectively by the indicated bias system ; but transient depressions of V_E after a pulse are inevitable because of the loss of charge on the coupling condenser between the first and last valve. They will change the bias and therefore shift the channel for the next pulse. Fortunately, this defect can be eliminated. For an input pulse of given height, point C always follows the input grid up to the same peak potential irrespective of the potential where it began to move. Consequently, point D too always attains the same peak potential. The discriminator should then measure this absolute potential rather than refer all measurements to the drifting waiting potential V_D (or V_E). Hence it is essential to d–c couple into the discriminator. Taking D as output terminal, putting $V_D=-70$ v (say) by a suitable choice of battery voltage, and adjusting the lower discriminator to trigger at -50 v, the condition $V_D=-(a+50)$ holds automatically because a is now variable according to $a=-50-V_D$.

Of course this d–c coupling will not eliminate errors due to pulse tails produced in the preceding main amplifier, nor can it avoid channel shifts due to loss of charge on the input condenser of the biased amplifier. If the pulse repetition rate n per sec becomes so large that the average spacing $1/n$ of successive pulses is comparable to (or even smaller than) the time constant on the input grid, the channel will shift upwards by approximately An volts, where A is the average area under a pulse in volts times seconds. This shift can be minimized by using a sufficiently small time constant, and avoided altogether by a recovery diode, but will not in general be serious because it falls under the heading of common drift.

§ 3. Particular Features

This kicksorter is designed to handle scintillation pulses from the fast phosphors commonly used in coincidence experiments, e.g. anthracene or, at the worst, NaI(Tl). Furthermore, it operates fast and delivers the

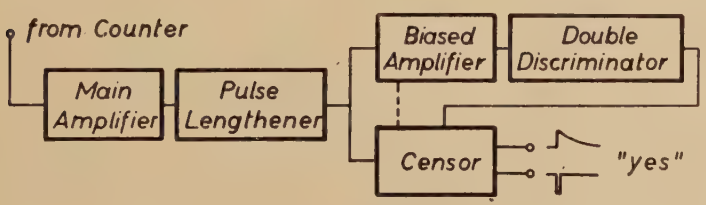
' yes ' in form of a 0·1 μsec-pip a well-defined and pulse height-independent time after the onset of the pulse so that it can be used in coincidence experiments with a resolving time of 0·1 μsec.

One cannot introduce such a message delay with precision if the pulse front has been slowed down by a high amplifier risetime. With a wide-band main amplifier thus necessary, it seemed easiest to give it only a small gain and achieve most of the amplification in a high-gain photomultiplier (EMI 6262 14-stage). These tubes show a considerable nonlinearity of response (cf. Greenblatt, Green, Davison and Morton 1952) which might sometimes be exploited to expand the low energy part of a spectrum at the cost of its upper end. To permit this, the main amplifier has variable gain and was designed with a view to good overload performance. It also accepts the pulses from the multiplier directly, without intermediate head amplifier, and barely produces any pulse tails at its output—a property very desirable in an amplifier preceding a kicksorter.

§ 4. Functional Arrangement

The block diagram of the complete instrument is shown in fig. 3. The pulse lengthening stage between main and biased amplifier gives the sorter enough time to operate, permits the use of bistable multivibrators in the double discriminator stage instead of self-resetting ones, and makes it

Fig. 3



Block scheme.

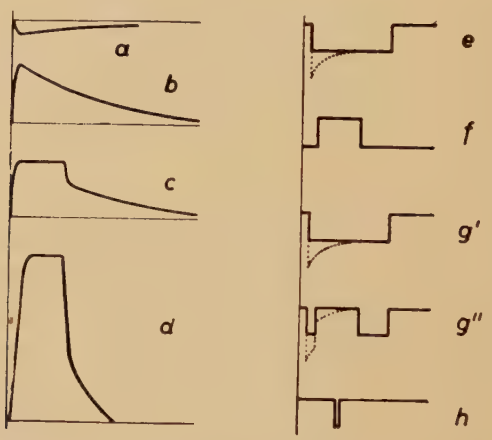
possible to collect the charges on the photo multiplier anode with a rather short time constant (2 μsec, say) which will effectively prevent falsifications of pulse height and shape by pile-up of several scintillation pulses or of a scintillation pulse and its ' satellites ' (Godfrey, Harrison and Keuffel 1951, 1952 ; Mueller, Best, Jackson and Singletary 1952). For the wave-forms of incoming, amplified and lengthened pulse, see figs. 4 (*a*), (*b*) and (*c*), respectively.

The biased amplifier output pulse, fig. 4 (*d*), produces current signals *e* and *f* in the lower and upper discriminator, respectively, which are added in a resistor to give voltage signal *g'* if the pulse falls within, or *g''* if it is higher than the channel. Distinction between the two is effected in a circuit called the censor. There a negative 0·1 μsec-pip is provided at a constant time after the pulse has come from the lengthener (fig. 4 (*h*)), and sent into coincidence with *g*. The coincidence mixer furnishes two output

signals : a negative ' yes ' congruent with the censor pip (for coincidence experiments) and a positive smeared-out ' yes ' at low impedance (for the channel count scaler).

To make the delay of the censor pip strictly independent of pulse height the censor should measure it from the moment at which the pulse has reached a given fraction (say half) of its total height. Of course this is impossible because the censor cannot know anything about the total height before it has been fully attained. Still, the delay of the ' yes ' will be to a sufficiently good approximation independent of pulse height if the censor is always alerted at the moment when the incoming pulse has reached half the *channel centre* height, and delivers the pip a fixed interval of time later : half the channel centre height corresponding very closely to half the total height of a pulse which falls into the channel, the censor pip will then arrive at the desired instant for such a pulse while for all others the exact spacing of the pip does not really matter because there will be no ' yes ' anyway.

Fig. 4



Waveforms. Length of time base ca. 3 μsec.

Thus the censor must be informed about the channel centre setting V_B, as indicated in fig. 3 by the dashed line, and is arranged to start operation whenever a pulse passes the threshold $|V_B|/2$. Incidentally, it follows that (*a*) the lower channel boundary must always be placed higher than half the channel centre height, (*b*) the delay of the ' yes ' depends on the rate of rise of the pulse, i.e. on the scintillator decay time.

§ 5. Circuit Diagrams

(i) *General*

The three supply lines carry +300 v, earth and −200 v, respectively. They are fed from feedback-stabilized standard power units which derive their standards of voltage from 85 Al neon stabilizers whose stability is stated by the manufacturers to be one part in 1000.

The following valves are used :

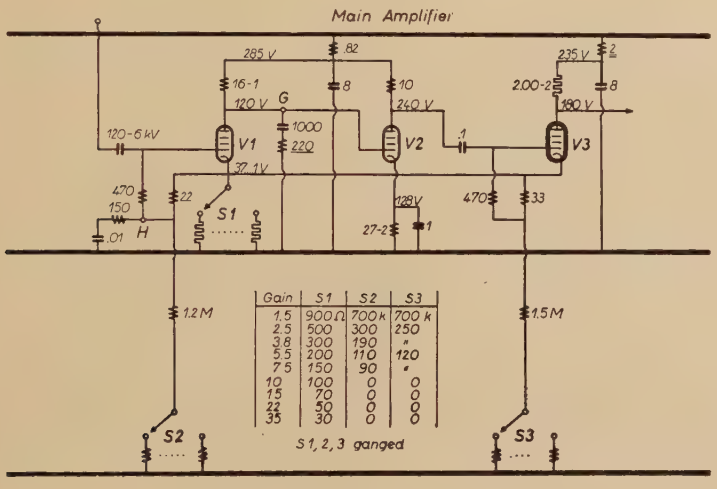
Crystal diodes	.. ..	CV 442
Diodes	.. .. ..	EA 50
Single triodes	.. ..	EC 91
Double triodes	.. ..	ECC 33
Pentodes	.. ..	EF 50, EF 55 (heavy symbol)

All screens go to anode supply (except where otherwise indicated), all suppressors to cathode.

Resistances are given in kΩ (except where otherwise stated). Values of high-stability carbon resistors are underlined once, of wire-wound resistors twice.

All capacitances lie between 12 pF and 8 μF and are stated in either pF or μF since no ambiguity can arise. Several non-essential decoupling condensers have been omitted in order to simplify the diagrams.

Reference potentials are derived from strongly decoupled points on wire-wound resistor chains across the supply lines (not shown).



Gain	S1	S2	S3
1.5	900 Ω	700 k	700 k
2.5	500	300	250
3.8	300	190	"
5.5	200	110	120
7.5	150	90	"
10	100	0	0
15	70	0	0
22	50	0	0
35	30	0	0

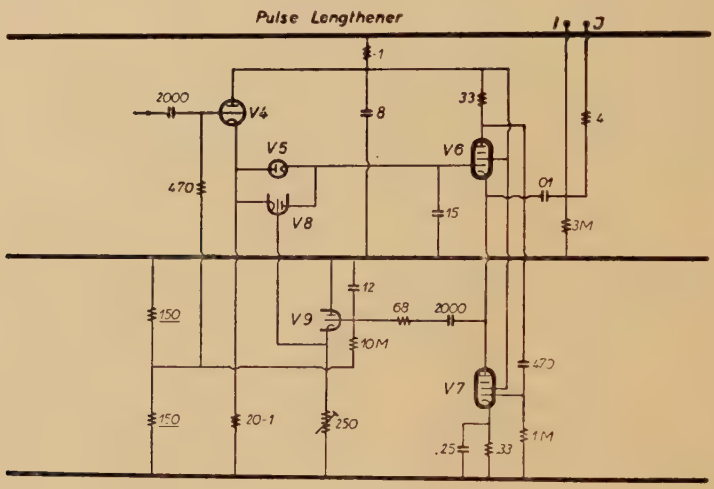
(ii) *Main Amplifier*

The main amplifier is a pulse-inverting ring of three with gain adjustable from 1·5 to 35. Its calculated bandwidth inside the loop is 20 Mc/s at the highest gain ; this figure is reduced to a gain-independent 10 Mc/s by the time constant at the anode of V3 where the pulse is taken off. The measured value is 8 Mc/s ; anthracene pulses are reproduced without harmful flattening.

V1 and V2 are d–c coupled to avoid excessive over-loading. A complicated bias system is then needed to keep d–c conditions throughout the amplifier the same for all gains. With the system used the input impedance is very high ; no negative pulse tails appear (except on over-loading, where the input impedance suddenly drops).

Critical compensation is achieved by the R–C phase-shift network from point G to ground (Scarrott 1951). One R–C combination suffices for all gains on account of the high loop gain (minimum 350, at the highest gain).

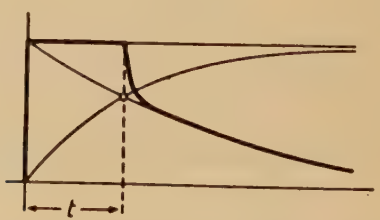
The small h.t. input condenser causes some trouble. Collaboration between the input and the various other time constants (which are all of the same order of magnitude) drives the loop into 1 kc/s oscillations at the higher gains. Since all these time constants advance the phase the situation is remedied by a R–C phase retarder from point H to earth, with a time constant well intermediate between that on the input and those of the decouplings.



(iii) *Pulse Lengthener*

Cathode follower V4 charges *via* V5 a memory capacity (15 pF + strays) connected to the input grid of White cathode follower V6 and 7. Across V5 lies the releasing valve V8. It is normally just cut off ; during a pulse

Fig. 5



Waveforms on V8 for ideally fast pulse. Anode : straight line (until further notice) ; cathode : exponential decay ; grid : exponential rise. Resultant pulse shape : heavy contour.

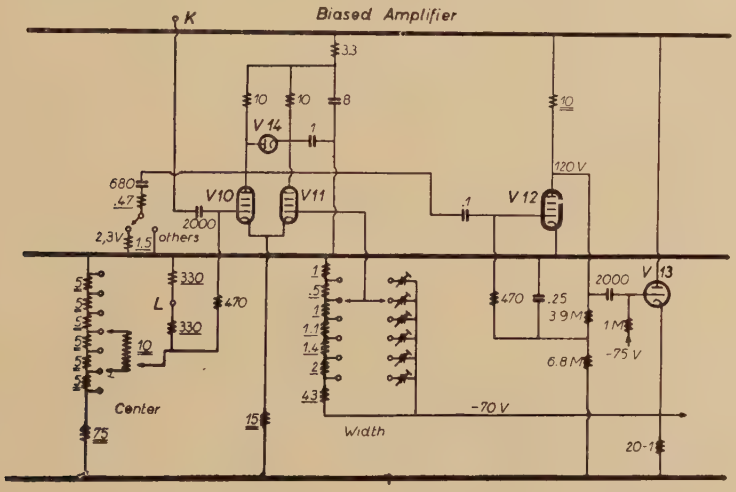
The releasing time t (ca. 1 μsec) does not depend on the ordinate scale, i.e. on pulse height.

its grid follows the exponential charging of a condenser (12 pF+strays) through a resistor (68 kΩ) to the full pulse height which is taken from the output of V6 and 7. The releasing action is explained in fig. 5. As V8 will inevitably draw grid current during release, cathode follower V9 is needed between the charging capacity and the grid of V8. V9 is adjusted to work at *low* g_m in order to reset V8 automatically near cutoff in case of any variation of d–c levels.

Low-impedance outputs I and J serve to drive coaxial cables to the biased amplifier and to a multi-channel kicksorter, respectively.

(iv) *Biased Amplifier*

Cathode follower V13 takes the place of the battery in fig. 2, providing low output impedance to tow the large stray capacities on the feedback attenuator and deal with grid current in the discriminator stage. The anode of V11 has to be steadied by a big condenser, otherwise positive feedback between V11 and V12, *via* V14, will cause violent oscillations.

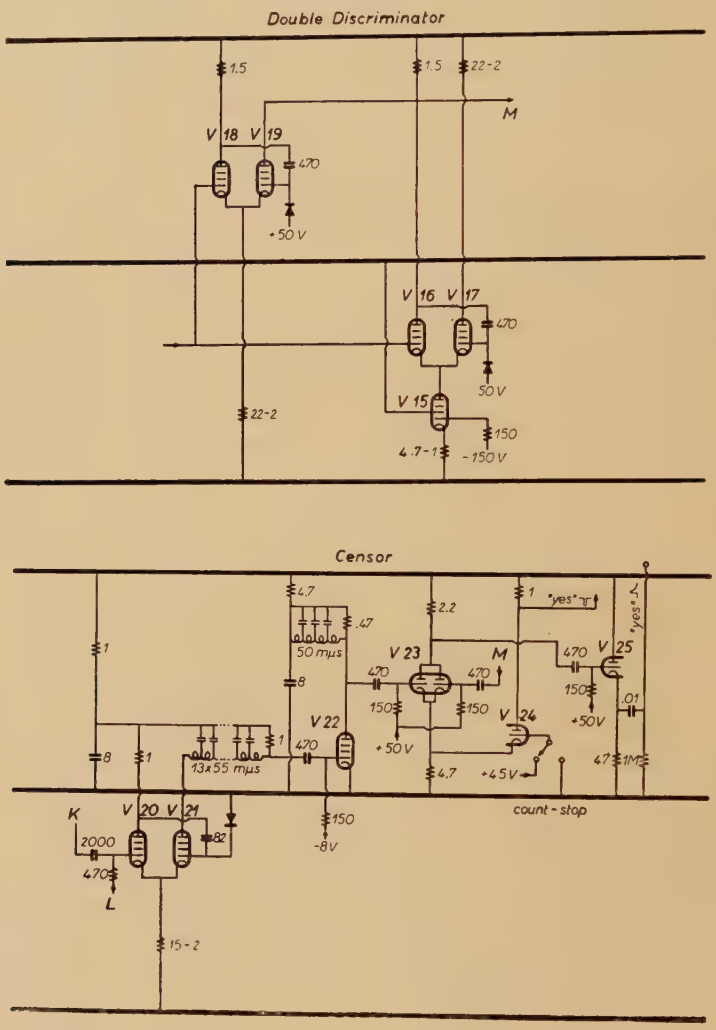


The gain is adjustable in steps corresponding to channels of width 2, 3, 5, 7, 10 and 15 v. The loop-gain is only about 50 for the highest gain, i.e. the 2v-channel, and since it is proportionately higher for the wider channels one R–C phase-corrector is here unable to achieve critical compensation at all gains (Scarrott 1951). The compromise of two corrector combinations (680 pF and 2 kΩ for the 2v and 3v channels, 680 pF and 0·5 kΩ for the others) has been adopted at the cost of a somewhat over-critical compensation for the narrower channels.

(v) *Double Discriminator*

The lower boundary-pair V16 and 17 needs constant-current cathode resistor V15 on account of the high grid swing. To maintain the reference potentials germanium recovery diodes are used. Their backward resistance is only 200 kΩ so that no parallel resistors are necessary.

The well-known ' current decay effect ' in valves waiting at cutoff (Eisenstein 1951)—here V16—makes the output pulse deviate from its ideal form as indicated in figs. 4 (*e*), (*g'*) and (*g''*) by the dotted contours. This phenomenon incidentally puts a lower limit on both the useful biased amplifier risetime and the delay of the censor pip.



(vi) *Censor*

Bistable trigger pair V20 and 21 is switched over by any pulse higher than $|V_B|/2$ and sends a positive pulse over a high-bandwidth 0·7 μsec-delayline into V22, where the censor pip is produced by means of a shorted 0·05 μsec-delayline.

The positive ' yes ' comes *via* V25 from the anode of coincidence valve V23, the negative ' yes ' from the anode of stopping triode V24 where, if

wanted, a delayline could be introduced for delayed coincidence experiments or to synchronize two kicksorters working on pulses from phosphors with different decay times. The grid of V24 provides a convenient place for a counting switch that controls positive and negative ' yes ' strictly simultaneously.

§ 6. Accuracy of Channel Definition

(i) *Relative Drift*

Superimposed on the 50 mv-uncertainties in the two discriminator levels are 50 mv-fluctuations in each of the reference potentials -50 v and $+50$ v, owing to the one part in 1000 stability of the voltage supplies. Promoting these four figures to the rank of r.m.s. standard deviations—a safe procedure because they are maximum variations really—we get a composite standard deviation of 100 mv or 0·1% of the 100 v-gap. Concerning g we estimate a maximum 10 hour variation of 5% in the slopes of V10 and V12, which gives $5\sqrt{2}$% variation in the gain without feedback, and divide by the loopgain. The latter is at least 50, so that the variation of g is at most 0·15%. The instability of the gain-defining resistors (R_1 and R_2 of fig. 2) will be much smaller than 0·15% and can therefore be neglected. Hence the relative standard deviation of $w=100/g$ should be less than 0·2% for all channels, permitting the full statistical significance of counts of 10^5 to be realized.

No bench test of this channel constancy was attempted because of the extreme difficulties in the way of producing a white spectrum of randomly spaced test pulses with a mean pulse frequency constant to better than 0·2% over 10 hours.

(ii) *Common Drift*

There will be 50 mv uncertainty in the biased amplifier effective input grid potential, combined with a 50 mv-fluctuation of V_B at the highest channel centre setting. Fluctuations of V_C are of no concern because compensated for automatically by the d–c coupling. The result is a r.m.s. common drift of at most 70 mv.

Relative drift also contributes to common drift. The two channel boundaries do not fluctuate symmetrically about an invariable channel centre, but move independently of each other. Therefore the instantaneous channel centre fluctuates about the mean centre even when there is no drift in bias. In the present circuit, these centre fluctuations are partly correlated with the fluctuations in V_B because both V_B and the lower discriminator reference potential are derived from the same supply voltage. As a little calculation shows, the net effect is then even a small decrease in the above 70 mv.

To the first order, gain variations in the main amplifier stage will also simulate common drift. However, in the present case they are negligibly small (neglecting instabilities of the gain-defining resistors in the anode and cathode of V3, the 10 hour variation of gain is at most $5\sqrt{3}$% divided by 350, i.e. 0·03%, or 15 mv in the highest pulse of 50 v).

Hum and drifts in the d–c output levels of main amplifier and pulse lengthener are differentiated away by the a–c coupling into the biased amplifier and cannot therefore shift the channel.

§ 7. Test Pulses

Testing of these circuits had to be carried out with test pulses of the exact shape of anthracene or NaI pulses (which reach their top after 0·15 and 0·6 μsec, respectively, for a 2 μsec pulse back). A simple pulse generator answering this description has been permanently included in the final setup of two kicksorters plus coincidence stage and coincidence scaler because the test pulses offer also a convenient means of comparing channel settings with the complete pulse spectra observed in a multi-channel kicksorter.

Acknowledgments

I wish to express my warmest thanks to Mr. G. G. Scarrott. His competent advice and efficient collaboration were of invaluable aid to me.

This work was carried out while the author held a British Council Scholarship.

References

Eisenstein, A., 1951, *J. Appl. Phys.*, **22,** 138.
Eppstein, J. S., 1951, *J. Sci. Instrum.*, **28,** 41.
Farley, F. J. M., 1953, *J. Sci. Instrum.*, **30**, in the press.
Francis, J. E., Bell, P. R., and Gundlach, J. C., 1951, *Rev. Sci. Instrum.*, **22,** 133.
Freunlich, H. F., Hincks, E. P., and Ozeroff, W. J., 1947, *Rev. Sci. Instrum.*, **18,** 90.
Gray, W. J., 1948, Appendix C in Valley and Wallmann (1948). (See below.)
Greenblatt, M. H., Green, M. W., Davison, P. W., and Morton, G. A., 1952, *Nucleonics*, **10,** No. 8, 44.
Godfrey, T. N. K., Harrison, F. B., and Keuffel, J. W., 1951, *Phys. Rev.*, **84,** 1248 ; 1952, *Nucleonics*, **10,** No. 3, 33.
Hutchinson, G. W., and Scarrott, G. G., 1951, *Phil. Mag.*, **42,** 792.
Johnstone, C. W., 1953, *Nucleonics*, **11,** No. 1, 36.
Mueller, D. W., Best, G., Jackson, J., and Singletary, J., 1952, *Nucleonics*, **10,** No. 6, 53.
Roulston, K. I., 1950, *Nucleonics*, **7,** No. 3, 27.
Scarrott, G. G., 1950, *Progr. Nucl. Phys.*, **1,** p. 89 ; 1951, unpublished calculations, to appear in Vol. 3 of *Progress in Nuclear Physics.*
Valley, G. E., and Wallmann, H., 1948, *Vacuum Tube Amplifiers*, Radiation Laboratory Series Vol. 18 (McGraw-Hill), p. 422.
Van Rennes, A. B., 1952, *Nucleonics*, **10,** No. 8, 22.
Wilkinson, D. H., 1949, *Proc. Camb. Phil. Soc.*, **46,** 508.

CIV. *Studies of Continuous Phase Transformations with* x-*Rays*

By E. Gillam and D. G. Cole
British Iron and Steel Research Association, Battersea, London, S.W.11*

[Received May 2, revised June 3, 1953]

Abstract

A method due to Heal and Savage (1949) and described in detail by Heal and Mykura (1950) has been developed to study the progress of the transformation of austenite (*a*) into martensite as the temperature falls below the Ms point, and (*b*) under isothermal conditions at temperatures below the Ms point. The martensite results do not agree entirely with an empirical formula of Cohen (1951); the isothermal results show how these transformations are changed markedly by the addition of 1 or 2% of alloying elements. The two sets of results are combined to draw the 'S-curves' below the Ms point for the steels EN24, EN26 and EN30B.

§ 1. Introduction

The physical properties of a steel depend markedly on the heat treatment to which it is subjected on cooling it from the austenitic (face-centred-cubic) state. A summary of how a given steel reacts when quenched to a given temperature, where it is subsequently held, can be conveniently presented as an isothermal transformation diagram, and atlases of these diagrams (also frequently called 'S-curves' or TTT diagrams) have been published in Britain (Iron and Steel Institute 1949) and in the U.S.A. (United States Steel Co. 1951). The methods used to obtain the results from which these atlases were compiled were dilatometric with a microscopic check, and the diagrams contain no data below the Ms point, for in this region reliable measurements are difficult to obtain. The Ms point is the temperature at which austenite first transforms into martensite during a sudden cooling from its stabilizing temperature, which is usually above 800°C. Cohen (1951) has obtained isothermal transformation data below the Ms point for certain steels using a metallographic method due to Greninger and Troiano (1940), and also gives an empirical formula for the percentage of austenite transformed into martensite before the onset of any subsequent isothermal reaction. These seem to be the only figures at present available.

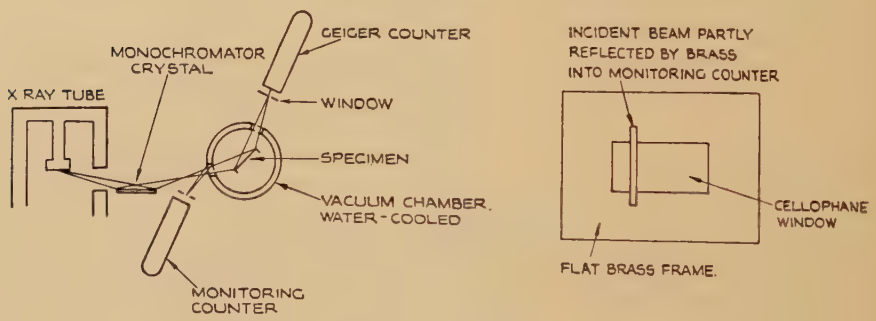
* Communicated by the Authors.

An x-ray method of studying continuous transformations has been described by Heal and Mykura (1950) who showed that for isothermal transformations above the Ms point the method gave results in good agreement with those done by the microscopic method. This paper describes how this method has been modified to give data both for the martensitic reaction down to a given temperature below the Ms point and for the subsequent isothermal reaction, from which the S-curves below Ms for the three steels EN24, EN26 and EN30B have been completed.

§ 2. Principle of Method and Apparatus

The principle of the method is to measure with a Geiger-counter the intensity of the diffraction line of monochromatic x-rays from the (220) planes of the austenite phase continuously as the transformation proceeds, the counting-rate at any instant being a measure of the amount of austenite present. For a transformation proceeding quickly, a counting-

Fig. 1



Plan of apparatus.

rate meter and a recorder can be used, and for one proceeding slowly monitored counts can be taken at convenient time intervals ; both of these methods have been employed. The steel specimens are of rolled strip 0·005 in. thick, and are mounted horizontally with the rolling plane vertical in a water-cooled vacuum chamber at a pressure of about 0·01 microns of Hg fitted with cellophane windows through which the x-rays enter and leave the chamber. The specimen is heated by passing through it a current of about 25 amp, and quenched by reducing the heating current so that the strip cools by radiation to the case ; a typical cooling rate so obtained is from 850°c to 300°c in 20 seconds. A chromel-alumel thermocouple spot-welded to the specimen records its temperature. The details of the apparatus used (shown in fig. 1) were very similar to those described by Heal and Mykura except for changes in the monitoring system which will be given later.

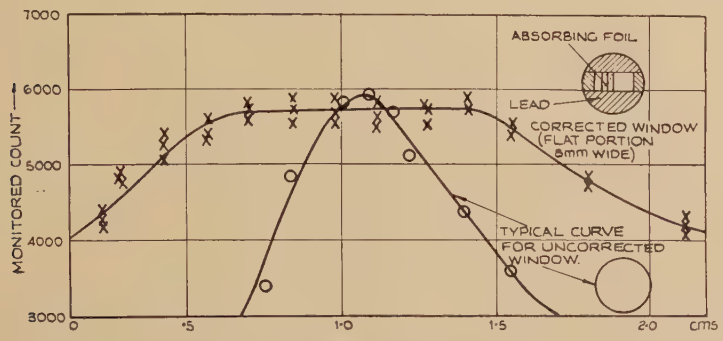
§ 3. Application of the Method

3.1. *The Martensitic Transformations*

The austenite-to-martensite transformation first occurs when the temperature reaches the Ms point, the value of which is characteristic for each steel, and on further cooling the transformation continues, the amount of austenite transformed depending on the temperature finally reached. To ensure that the specimen is 100% austenitic at the Ms point the quench should be rapid enough to avoid any other phase changes above the Ms point ; for the three steels examined this required a quench from 850°C to 250°C in less than 45 seconds. Cooling by radiation was considerably faster than this, as noted above.

To follow the course of the transformation to martensite one must measure the integrated intensity of an austenite line while the temperature changes. Due to thermal contraction the lattice parameter of the austenite and hence the Bragg diffraction angle changes with temperature,

Fig. 2



Variation of counter sensitivity across counter window (measured as variation with Bragg angle).

obeying the expression $d\theta = \alpha \tan \theta dT$, where $d\theta$ is the change of Bragg angle for a change of temperature dT and α is the coefficient of linear expansion. Thus it was found that during a quench of 700 degrees the angular displacement of the diffracted beam corresponded to a movement across the Geiger-counter window of 2·5 mm. It was necessary therefore that the counter should have uniform sensitivity across its window to accommodate this movement. The Geiger-counter* used had a circular end-window of mica and showed a sharp peak in its sensitivity curve (fig. 2) ; this was made flat as shown in fig. 2 by placing sections of lead to form a wide rectangular slit in front of the window and pieces of absorbing foil across this where necessary.

The temperature dependence of the x-rays diffracted by a crystal lattice had also to be considered. The integrated intensity, I, of the diffracted beam from a mosaic crystal is proportional to

$$\exp\left[-(2B \sin^2 \theta)/\lambda^2\right],$$

* Made for x-ray work by G. E. C. (Type XAI).

where $\exp[-(B\sin^2\theta)/\lambda^2]$ is called the Debye–Waller temperature factor. In this, λ is the wavelength of the monochromatic x-rays used, θ is the Bragg Angle for the planes considered, and

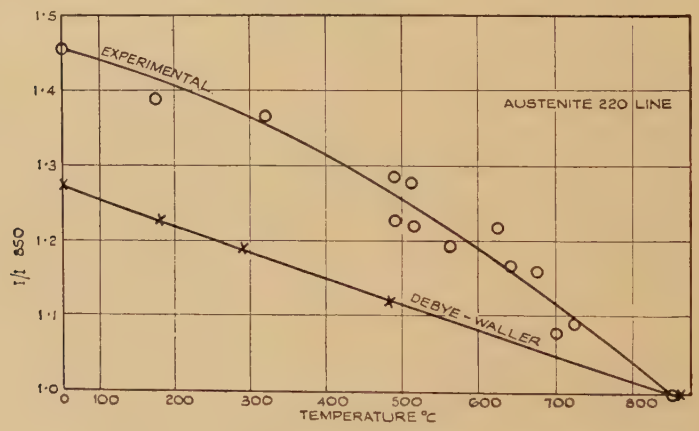
$$B=\frac{6h^2}{mk\Theta}\left(\frac{1}{4}+\frac{\phi(x)}{x}\right),$$

where $x=\Theta/T$ and $\phi(x)=\frac{1}{x}\int_0^x \frac{\xi d\xi}{e^{\xi}-1}$ (see James 1948, chapter 5).

Here h is Planck's constant, k is Boltzmann's constant, Θ is the Debye characteristic temperature and m is the atomic mass, T being in degrees absolute.

Figure 3 shows the intensity variation with temperature calculated from this formula and that given by experiments on an austenitic steel

Fig. 3



Variation of diffracted x-ray intensity with temperature.

using the modified counter described above. It is seen that the Debye–Waller theoretical factor is lower than that found in practice. In the present experiments the austenite (220) line was used; Birks (1950) found similar deviations for the (III) line. James (1948, pp. 231–233), discussing his own and other workers' data has also observed that the Debye–Waller formula does not hold at high temperatures.

However, the experimental results show that the intensity variation factor in the range of the martensite transformation, i.e. between 300°C and 150°C is still small and neglecting it altogether makes a difference of only between $\frac{1}{2}\%$ and 1% in the calculated percentage transformed.

The martensite studies, then, were carried out as follows: after austenitizing, the specimen was quenched and synchronous measurements recorded of the specimen temperature and the counting rate. The Ms point was reached in about 20 seconds and the whole operation completed

in 4–5 minutes. For such short runs the x-ray intensity was considered sufficiently stable ; any run in which fluctuations of the tube voltage and current were observed was rejected. Another experimental precaution taken was that the x-ray beam fell on to that part of the specimen possessing the spot-welded thermocouple.

3.2. *Application to the Isothermal Transformations below* Ms

If the austenite–martensite transformation is halted by stopping the quench at a temperature below the Ms point where it is subsequently maintained, an isothermal transformation of any remaining austenite occurs, though the onset of this is usually delayed and may take some minutes after the particular temperature is reached. As this is an isothermal phase change, the counter modification and the Debye–Waller factor noted in the last section need not be considered, but since these transformations can take several hours to be completed the instability of the x-ray tube output must necessarily be allowed for. A monitoring system was employed, each count measuring the intensity of austenite reflections being taken, not over a fixed time interval, but over a fixed number of counts on a second counter activated by a part of the incident beam reflected by the brass vacuum-chamber near the cellophane window. Monitoring systems have been described by Hall, Arndt and Smith (1949) Hargreaves, Prince and Wooster (1952) and Gillam and Heal (1952), but the system used was chosen for reasons given by Gillam and Cole (1952). A similar one has been described recently by Coyle, Hale and Wainwright (1953). For these more lengthy transformations the practice was to take monitored counts at intervals of one minute after the particular temperature being considered had been reached, and at longer intervals where the reaction was very slow, the counts being plotted on a logarithmic time scale.

§ 4. Some Considerations of Specimen Treatment

To conform to conditions specified in the atlas of the Iron and Steel Institute, the three steels should be austenitized for 30 minutes at their quoted temperatures before each quench. This period is to ensure that any carbides present before austenitizing are dissolved and to allow a stable grain size to be reached. Under these conditions it was found that noticeable decarburization occurred after five or six experiments on the same steel with a pressure of 0·1 micron of mercury. Consideration was given to the suitability of using a shorter annealing period for the specimens, as data obtained by Cole, Feltham and Gillam (1953) on grain growth in austenite showed that a relatively stable grain size was obtained in times as short as 10 minutes after the nucleation of the austenite phase. The grain size measurements in this work were done using thermal etching, the specimens and apparatus being identical with those used in the present method. Comparisons were then made between the progress of transformations after 30 and 10-minute annealing times. For an

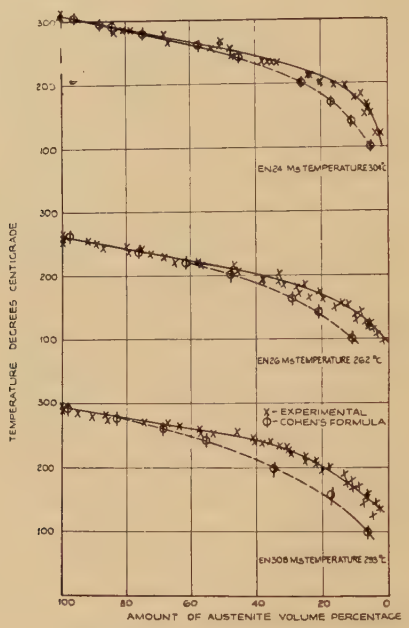
isothermal above the Ms point the results could also be compared with the published S-curves. From these comparisons it was found that there were no systematic deviations,

(*a*) in successive runs on the same specimen with 10-minute annealing, and

(*b*) in runs with a 30-minute anneal compared with runs with a 10-minute anneal

and that all cases showed good agreement with the published curves above the Ms point.

Fig. 4



Martensitic transformations.

Chemical analysis showed that up to six successive runs could be made with one specimen using 10-minute heat treatment without noticeable loss of carbon, and so most subsequent results were obtained with 10-minute austenitizing restricted to six runs per specimen. Occasionally a single 30-minute run was included which never differed significantly from the main results.

§ 5. Results on the Martensitic Transformation

The amount of austenite remaining untransformed at a temperature T°C, which is below the Ms point, on cooling from the austenitic state is given in fig. 4 for the three steels EN24, EN26 and EN30B whose compositions are given in table 1.

Table 1. Specimen Composition (by weight)

Steel	%C	%Si	%Mn	%S	%P	%Ni	%Cr	%Mo	
EN24	0·38	0·20	0·69	0·010	0·017	1·58	0·95	0·26	Fe to 100%
EN26	0·42	0·31	0·67	0·022	0·029	2·53	0·72	0·48	
EN30B	0·32	0·29	0·47	0·020	0·022	4·13	1·21	0·30	

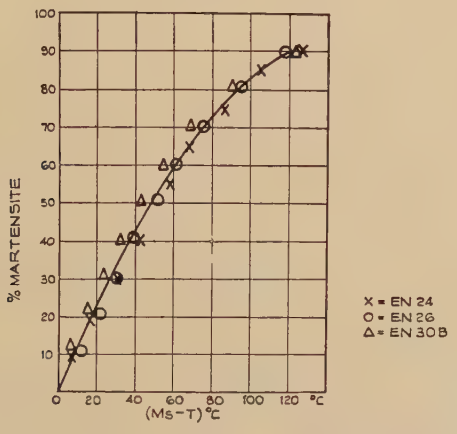
(EN30B also contains 0·11%W, 0·15%Cu, 0·013%V)

Reference has already been made to an empirical formula by Cohen (1951) giving the percentage of martensite P, formed instantaneously at a temperature Tq (in degrees c) prior to any subsequent isothermal phase changes :

$$P=100-1{\cdot}11\times10^{-12}(455-(Ms^*-Tq))^{5{\cdot}22}.$$

Ms* in this formula is not the temperature at which first traces of martensite appear but an extrapolated Ms temperature. Cohen claims

Fig. 5



Amount of martensite against temperature below Ms.

that this formula holds for steels containing 0·75 to 1·35% carbon, 0 to 2·8% chromium and 0 to 5·4% nickel, with austenitizing temperatures varying from 845°c to 1040°c, and is independent of chemical differences in these ranges, the presence or absence of undissolved carbides, variations in grain size and differences in quenching stresses by rapid cooling to different temperatures.

In the present x-ray method, an accurate value for the temperature at which the martensite transformation first begins is obtained since a very sharp drop in the intensity occurs at this point. Let this temperature be Ms. Then from Cohen's formula

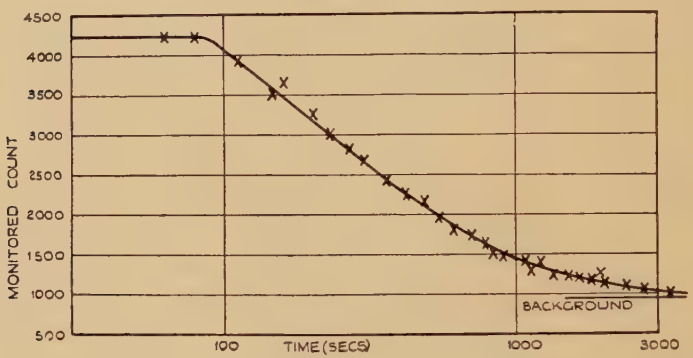
$$0=100-1{\cdot}11\times10^{-12}(455-(Ms^*-Ms))^{5{\cdot}22}$$

so that a knowledge of Ms for each steel gives a corresponding value

of Ms*. In this way Cohen's formula can be compared with the results of the present investigation, P then being calculated for any value of Tq. The comparison is made in fig. 4 where it is seen that this formula and the present experimental results agree closely in the first part of the transformation, but then differ as less austenite remains. This may be due to the lower amount of carbon in the present steels.

In fig. 5 the results are shown in a different way, with the percentage of austenite transformed plotted against the value (Ms–Tq). Although the Ms temperatures for the three steels differ by more than 40°C, all these points lie approximately on the same curve, showing that the significant variable on which the amount of martensite formed depends is the degree of cooling below the Ms point.

Fig. 6



A typical trace.

§ 6. The Results for Isothermal Transformation below the Ms Point

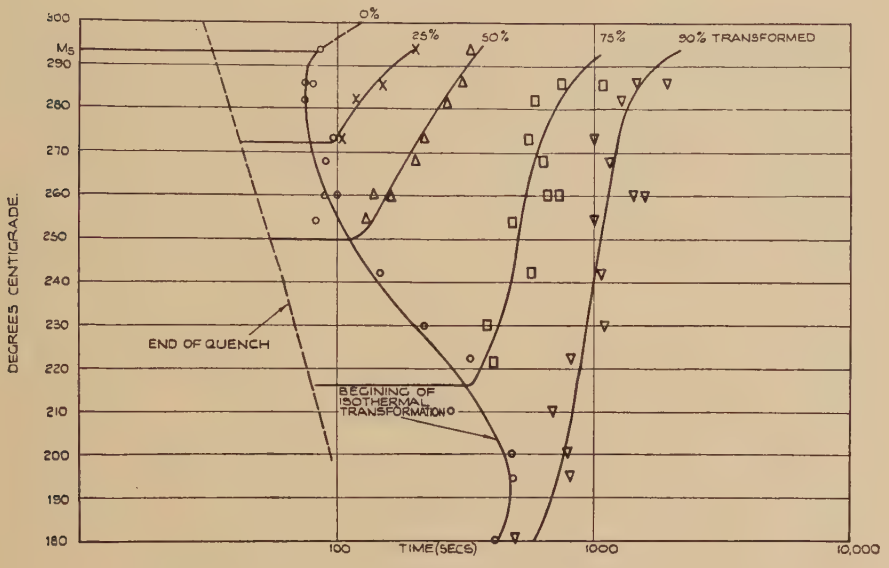
From the results already discussed, the percentage of austenite remaining for isothermal transformation at any temperature below the Ms point was known. For each such temperature the monitored count was plotted against time, measured from the beginning of the quench, on a logarithmic scale. This showed that after the isothermal transformation stage had been reached, the amount of austenite remained constant for a short period and then decreased linearly at first with log-time. A typical result is shown in fig. 6 where, it is noted that there is no sharp ending to the transformation. The count corresponding to zero austenite was taken by moving the Geiger-counter off the line into the background, or by leaving the specimen at the same temperature for periods well beyond the expected time for 100% to have transformed. On those occasions when both methods were used they showed close agreement.

For temperatures well below the Ms point where only a small percentage of austenite remained after quenching for isothermal reaction, a higher peak-to-background ratio was obtained by decreasing the width of the

slit in front of the measuring counter, and setting the counter slit on a previously determined position of the line at the temperature in question.

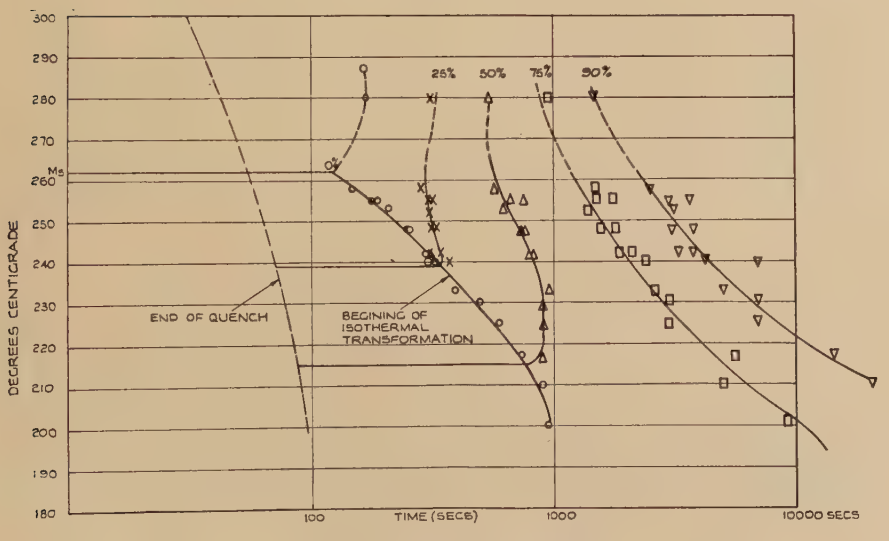
The ' S-curves ' below the Ms point for the three steels investigated are shown in figs. 7, 8 and 9. For the steel EN24 containing $1\frac{1}{2}\%$ of

Fig. 7



' S-curve ' below Ms for EN30B.

Fig. 8



' S-curve ' below Ms for EN26.

nickel and 1% of chromium the rate of isothermal transformation at temperatures immediately below Ms increases slightly for about 20° and then decreases as the temperature is reduced. The time for the

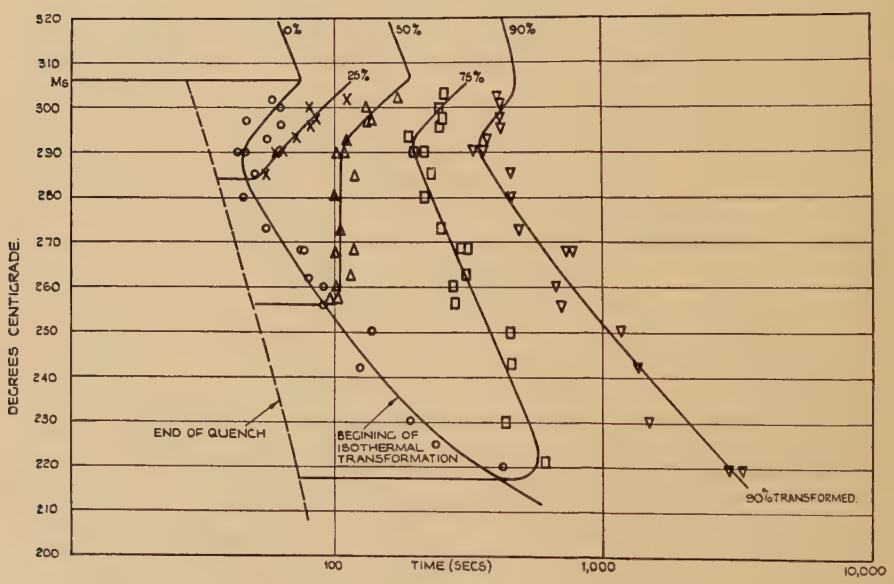
isothermal transformation to begin, shown by the dotted curve, shows a corresponding variation. Cohen (1951, p. 611) has also observed that the isothermal bainite reaction just below Ms is stimulated by the presence of coexisting martensite, but that this effect diminishes at lower temperatures.

The steel EN26, containing only an extra 1% of nickel, on the other hand, shows a reduced rate of transformation continuously as the temperature decreases below Ms, eventually becoming very slow indeed.

Finally with EN30B, which has a further 1½% of nickel and slightly more chromium together with some copper and tungsten, the transformations become progressively more rapid as the temperature decreases below Ms.

It is well known however, that much larger additions of alloying elements especially nickel, result in the virtual prevention of the trans-

Fig. 9



' S-curve ' below Ms for EN24.

formation, as in austenitic steels. This method of obtaining complete S-curve data would be well suited to a systematic study of the effect of alloying element on phase changes, using a series of carefully prepared specimens containing graduated amounts of alloying elements.

§ 7. Accuracy of the Method

This method has the usual limitations on accuracy found in work with Geiger-counters, and the effect on the accuracy of monitoring with a second counter is discussed by Gillam and Heal (1952). The accuracy

varied considerably in the present experiments since the total count which determines the error, differed widely in each of the measurements. This was partly because at low temperatures little austenite was left to transform, and partly because of the changes made in the slit width. An estimate of the error can be made from the scatter on the graphs displaying the results. For the martensite figures, the average curves are probably correct to ± 2 or 3% of martensite at any temperature; for the isothermal curve the error of the 75% and 90% lines may be $\pm 10\%$ though since these are normally plotted on a log scale it does not appear to be so serious.

A higher accuracy can always be obtained using Geiger-counter methods by increasing the intensity of the incident x-ray beam. In the experiments described, the tube voltage and current were 50 kilovolts and 10 milliamps respectively.

§ 8. Conclusions

The x-ray method of observing continuous phase transformations can be adapted to obtain information on the progress of (*a*) martensitic transformations, (*b*) isothermal transformations below the Ms temperature. Its ease and rapidity make it suitable for a systematic study of the effect of alloying elements on transformation rates. The results of the present experiments on the steels EN24, EN26 and EN30B have shown that the rate of the martensitic transformation during continuous cooling agrees well with an empirical formula of Cohen initially, but deviates from it as the transformation proceeds possibly because of the lower carbon contents of steel used in the present experiments. Also the amount of austenite transformed at any temperature T°C which is below Ms depends on the value (Ms–T) in the same way for all three steels although their Ms points differ by 40°C. Comparing the isothermal transformation data for EN24 and EN26 the effect of an additional 1% of nickel is to retard the isothermal reaction at lower holding temperatures; but the further addition of nickel, together with other alloys as in EN30B, re-stimulates it.

Acknowledgments

The authors wish to thank Mr. M. W. Thring and Mr. J. Savage at B.I.S.R.A. for their continued support and many useful discussions during the course of the work.

References

Birks, L. S., 1950, Naval Research Laboratory Report 3749, N.R.L., Washington, D.C.

Cohen, M., 1951, *Phase Transformations in Solids* (London: Chapman and Hall), p. 608.

Cole, D. G., Feltham, P., and Gillam, E., 1953, BISRA Report P/C.172.

Coyle, R. A., Hale, K. F., and Wainwright, C., 1953, *J. Sci. Instrum.*, **30**, 151.

Gillam, E., and Cole, D. G., 1952, *BISRA Report* P/N. 164.

Gillam, E., and Heal, H. T., 1952, *B. J. Appl. Phys.*, **3**, 353.

Greninger, A. B., and Troiano, A. R., 1940, *Trans. A.S.M.*, **128,** 537.
Hall, W. H., Arndt, U. W., and Smith, R. A., 1949, *Proc. Phys. Soc.* A, **62,** 631.
Hargreaves, C. M., Prince, E., and Wooster, W. A., 1952, *J. Sci. Instrum.*, **29,** 82.
Heal, H. T., and Mykura, H., 1950, *Metal Treatment*, **17,** 129.
Heal, H. T., and Savage, J., 1949, *Nature, Lond.*, **164,** 105.
Iron and Steel Institute, 1949, *Atlas of Isothermal Transformation Diagrams*, Special Report No. 40.
James, R. W., 1948, *The Optical Principles of the Diffraction of X-rays* (London : Bell).
United States Steel Company, 1951, *Atlas of Isothermal Transformation Diagrams.*

CV. *On the Annihilation Spectrum of Positrons in Freon and Oxygen*

By G. M. Lewis and A. T. G. Ferguson
Department of Natural Philosophy, The University, Glasgow*

[Received July 6, 1953]

Abstract

An examination of the radiation arising from the annihilation of positrons, from ^{22}Na, in the gases freon and oxygen, in delayed coincidence with the nuclear 1·28 mev γ-ray, has been made, using scintillation counters. The annihilation spectrum in freon, so obtained, has been compared with the theoretical three-quantum continuous spectrum of Ore and Powell, taking into account the characteristics of the spectrometer crystal ; and the curves are substantially similar. The spectrum in oxygen, on the other hand, indicates a single γ-line at 510 kev. Some decay times have been measured. The results are discussed.

§ 1. Introduction

The two-quantum annihilation of slowly moving positrons and electrons on Dirac theory (Dirac 1930) requires the system to be in a singlet state (cf. too Wheeler 1946). Ore and Powell (1949) have shown theoretically that annihilation from a triplet state is possible by three-quantum emission, and they have estimated a lifetime, and the form of the continuous spectrum. Deutsch (1951) has found a lifetime near this value using positrons in freon, and Deutsch and Dulit (1951) have shown the existence of radiations below 510 kev. De Benedetti and Siegel (1952) have shown the presence of three simultaneous annihilation quanta in freon. It seemed very desirable to examine the spectrum of this long period decay in freon, and to compare it with the predicted three-quantum continuous spectrum. The method of delayed coincidence appeared appropriate. This method, while selecting the required long period radiation, would also eliminate γ-radiations from the source and from the gas container. Positrons in oxygen have been shown by Deutsch (1951) to give a decay time inversely proportional to pressure as would be expected on simple two-quantum theory. It seemed of value therefore to investigate the spectrum in oxygen also, using the delayed coincidence method.

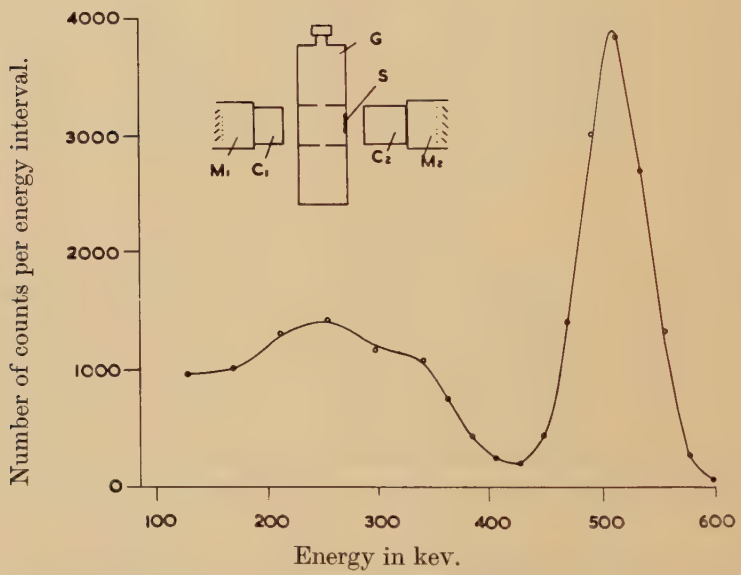
§ 2. Experimental Arrangement

A suitable source of positrons is ^{22}Na ; we have shown the delay between the emission of positrons and the 1·28 mev γ-ray to be less than 10^{-9} sec. A thin ^{22}Na source of strength $\sim 2\,\mu$c was mounted on and

* Communicated by Professor P. I. Dee.

covered by aluminium foil, 0·005 in. thick. It was placed as shown in fig. 1 inset, inside a cylindrical gas container 2 in. diameter, the wall of which was of 0·012 in. thick copper. Two perforated phosphor bronze plates 0·005 in. thick, $1\frac{1}{2}$ in. apart were used to keep the positrons in the central region. The counter nearer the source detecting the nuclear γ-ray had terphenyl in toluene as scintillator, occupying a volume $1\frac{1}{2}$ in. diameter $\times$ 2 in. (C_2). The other counter, opposite, acting as spectrometer for the annihilation radiation, had a block of sodium iodide, $1\frac{1}{4}$ in. cube (C_1). This block was composed of two pieces cemented together along a common plane with silicon grease, as no single crystal of suitable size was available. The scintillators were packed in magnesium

Fig. 1



Pulse height distribution obtained for the annihilation radiation of ^{64}Cu positrons in copper.
Inset: experimental arrangement; S, source; G, gas container; C_1, C_2, scintillators; M_1, M_2, multipliers.

oxide, and mounted on 14-stage E.M.I. tubes, type 6262. The ends of the iodide and terphenyl scintillators were situated 1 in. and $\frac{1}{2}$ in. respectively from the container. The system was set up on a light plywood table to reduce scattering.

The electronic arrangement is shown schematically in fig. 2 and involves a fast coincidence unit working essentially in the same manner as that described by Bell, Graham and Petch (1952). Pulses from the two counters can saturate their respective pre-amplifiers which deliver pulses through variable delaying cables of Uniradio 31 to a shorted line. The side channel, delivering attenuated pulses from the terphenyl counter, permitted the passing of pulses from γ-rays above 600 kev only; i.e.

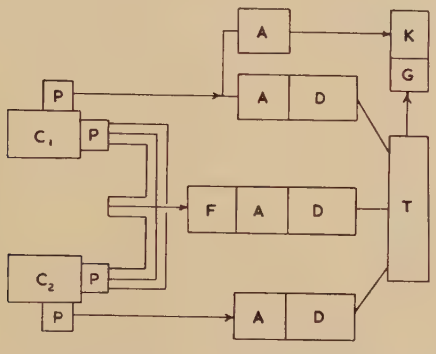
from the 1·28 MEV γ-rays only, here. On the sodium iodide side channel, pulses were taken from the tenth dynode to avoid saturation effects and the discriminator was set so that pulses from the collector were rapidly rising. The output from the triple coincidence unit opened the gate of the kicksorter analysing the pulses from the iodide crystal. The resolving time of the coincidence unit in this application was 5×10^{-8} sec. It may be mentioned in passing that a similar unit has been operated here with terphenyl counters using E.M.I. tubes with a resolving time of 10^{-9} sec.

The operation of the kicksorter depended on the display of brightened pulses on a cathode ray tube, which could be scanned by a multiplier, or photographed.

§ 3. Delay Times

With the container evacuated, no real coincidences occur with more than 20 metres of delay cable present on the 1·28 MEV γ-side (1 metre of cable was equivalent to 5×10^{-9} sec). For the determination of delayed spectra the unit should have resolving times roughly independent of the energies of the pulses from the iodide crystal. Figure 3 (*a*) shows the variation in coincidence counts with delay cable length, for a narrow energy

Fig. 2



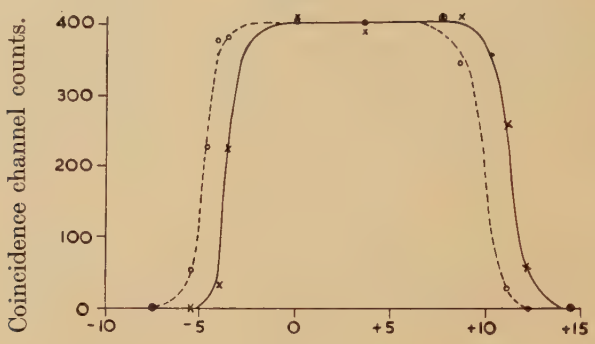
Block diagram : C_1, C_2, counters ; P, preamplifier ; F, unit receiving fast coincidences ; A, amplifier ; D, discriminator ; T, triple coincidence unit delivering selected fast coincidences ; G, gate ; K, kicksorter.

band of pulses from the iodide crystal, appearing in one channel of the kicksorter. It is seen that the resolving time is approximately independent of pulse height. There is however, a displacement of the equivalent of $1\frac{1}{2}$ metres of cable, viz. $7{\cdot}5\times10^{-9}$ sec, for the 200–400 kev change involved, the larger pulses arriving faster. This time is small compared with the decay times in the gases, as is seen below.

Figure 3 (*b*) shows the delayed coincidence curve when freon is present at a pressure of 4·8 atmospheres, approximately the pressure used when the spectra were determined. The lifetime from this and similar curves was $(1{\cdot}25\pm0{\cdot}10)\times10^{-7}$ sec. This time is only a little shorter than that obtained by Deutsch (1951) at lower pressures ; confirming further that

the variation of time with pressure is, at most, slight. The lifetime in oxygen at 2·4 atmospheres, the pressure used later, has been determined by Deutsch at 8×10^{-8} sec, and we have confirmed this value approximately.

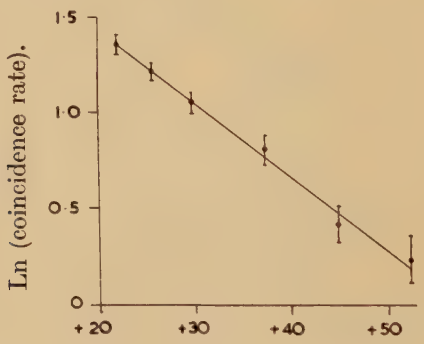
Fig. 3



(*a*)

Delay cable length in metres.

Coincidence resolution curve between prompt pulses of a given height in the sodium iodide counter and 1·2 Mev γ-pulses in the other counter.
x——electron energy, 200 kev; o– – –electron energy, 400 kev.



(*b*)

Delay cable length in metres.

Variation of coincidence rate with delay for positrons in freon at 4·8 atmospheres.

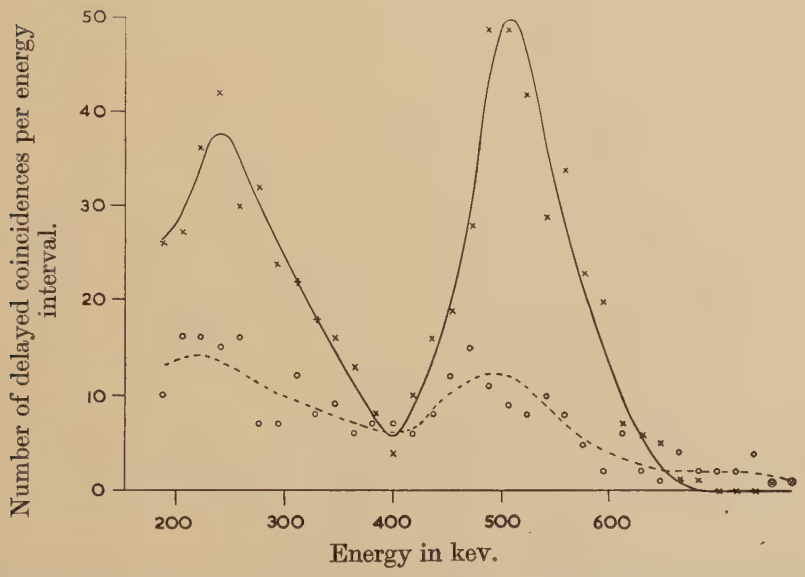
§ 4. Spectroscopy

It was first necessary to determine the spectrometer characteristics of the iodide crystal. Figure 1 shows the spectrum obtained when ^{64}Cu positrons were allowed to annihilate in copper, when only the iodide channel and kicksorter were employed (the container was not used in this run). The 1·38 Mev γ-ray of ^{64}Cu is only of $1\frac{1}{2}\%$ relative intensity (cf. Kubitschek 1950) and is not of importance here. The small low energy

bulge (above ~1000 counts) in the multiple Compton distribution can be explained by back scattering of γ-rays from the rear parts of the crystal mounting. A similar curve was obtained from ^{22}Na in the evacuated container with the two counters in prompt coincidence, when the 1·28 MEV γ-counter was positioned to minimize the prompt coincidences from its back scattered Compton radiation.

The effect which the position of the positron, at annihilation, would have on the observed spectrum, has been examined. Curves of the fig. 1 type have been taken with the ^{64}Cu source enclosed in copper, situated at various points in the region 1 in.–3 in. from the iodide crystal, also with ^{203}Hg (280 kev γ-ray) and ^{137}Cs (660 kev γ-ray) sources (screened when necessary from conversion x-rays). These spectral curves were little affected by positioning, the variation in the ratio peak counts to total counts being less than ~5%.

Fig. 4



Pulse height distribution obtained for the delayed annihilation radiation of positrons in oxygen.
x——real counts, after subtracting randoms; o– – –random counts.

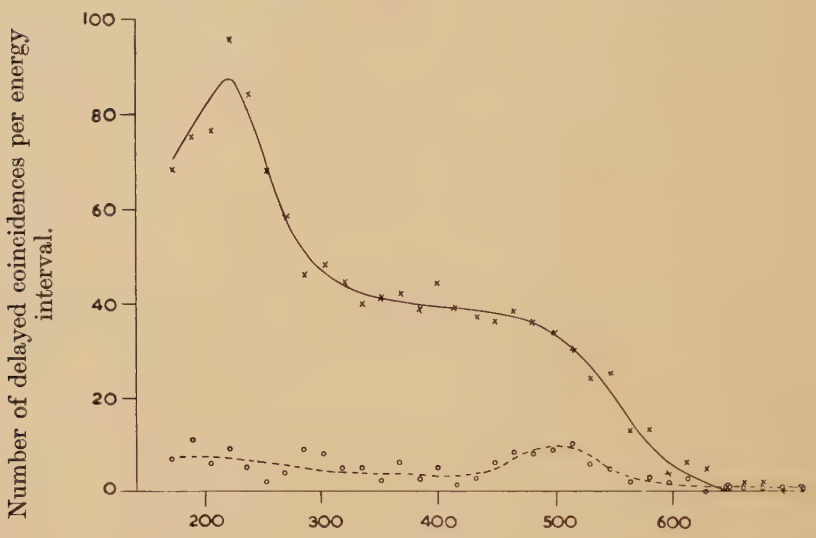
In the delayed coincidence spectrum experiments, the 1·28 MEV γ-channel was delayed by 25 metres of cable (cf. fig. 3 (*b*), the point +25). The discriminator on the iodide channel was set at ~140 kev to ensure satisfactory operation of the fast coincidence unit, and the pulses photographed. Pulses clear of the threshold were recorded. The spectra of the random coincidences were obtained by inserting the 25 metres of cable on the iodide channel (the point −25 on fig. 3 (*a*) produced back).

Figure 4 shows results for oxygen at 2·4 atmospheres. The dotted curve shows the random counts. The full curve shows the real counts

after subtracting the random counts, point for point. The presence of a main annihilation peak at 510 kev is seen, together with a low energy bulge corresponding to that in fig. 1 for copper. The bulge is more prominent in height, the reasons for this are given in § 5 below.

Figure 5 shows results for freon at 5·2 atmospheres. Again the full curve shows the real counts, after the subtraction of random counts. Here the 510 kev peak is absent, and the broad plateau exhibits clearly the inhomogeneous character of the annihilation radiations. These matters are discussed in detail below.

Fig. 5



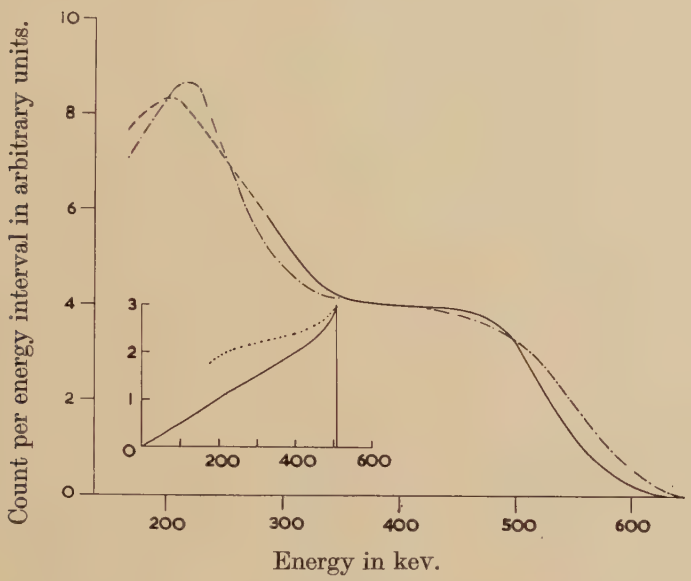
Pulse height distribution for the delayed annihilation radiation of positrons in freon.
x——real counts, after subtracting randoms ; o– – –random counts.

§ 5. Analysis of the Results

In the case of oxygen, the main peak at 510 kev is less pronounced in height compared with the low energy bulge than it is in fig. 1. This is due in great measure to the larger width at half height of this main peak (20%, compared with the 14% of fig. 1). This increased width is due mainly to a variation in gain over the several hours of running time involved in the measurement, and to a lesser extent to a slight gradual deterioration in the quality of the crystal since the experiments first began. Further a 10% increase is expected in the low energy pulses due to the displacement of the resolution curve (cf. § 3). Lastly measurements on scattering would suggest a rise $\sim$5% at the lower energies ; particularly due to the proximity of the terphenyl counter. It seems clear therefore that within the limits of the experiment, over the range investigated, the annihilation radiation consists of a single line, at 510 kev.

The case of freon is different. The theoretical Ore–Powell curve is shown in the inset of fig. 6. The lower energy quanta would be preferentially absorbed by the iodide crystal. Furthermore this preferential absorption is slightly accentuated by the proximity of the source points to the crystal, over the energy range involved. Using the γ-ray absorption data given by Heitler (1944), the dotted inset curve of fig. 6 shows the relative number of annihilation γ-rays of the various energies in the range considered, which would in fact be absorbed by the crystal.

Fig. 6



Main curves——— Pulse height distribution expected ;
.—.—.—. Experimental curve of fig. 5.
Inset :———Ore–Powell theoretical curve for three-quantum annihilation ;
..... expected photon response with particular NaI crystal.

Finally, the lower energy γ-rays absorbed in the sodium iodide crystal are more efficient in producing electrons in the peak essentially because the photoelectric effect becomes increasingly important compared with the Compton effect. Experimentally the ratio of electrons in the peak to the total number energized were 0·35, 0·43, 0·65, for the 680, 510, 280 kev γ-rays investigated, respectively. The expected electron spectrum not corrected for varying peak widths is shown by the main curve of fig. 6. The lower energy portion, shown by dashes, involved slight extrapolation but this part of the curve is essentially correct in shape and the errors there are estimated at less than $\sim$5%. The possibility of two quanta registering simultaneously has not been introduced here ; for most of the radiations would arise from positrons near the source, both for freon and oxygen ; and further two quanta running close together are not favoured in the theory, on phase space grounds.

The experimental curve of fig. 5 is also shown, for comparison, in fig. 6 ; and is seen to be in close agreement with the theoretical expectations over the range investigated.

§ 6. Discussion

There are general grounds for believing that the long period decay observed in freon is associated with only triplet bound states. For annihilation of the latter is a third order process, unlike that of bound singlets ; furthermore at higher pressures annihilation of free positrons, involving two and three quantum processes in competition, would be very fast. Again the small pressure dependence of the decay constant indicates that few triplets are converted to singlets by collision.

The proximity of the experimental value, in freon, of the decay time to the theoretical Ore–Powell triplet time of $1{\cdot}4\times10^{-7}$ sec, and the closeness of the experimental and theoretical spectra associated with these, confirm the nature of the state and afford verification of the theory. It should be mentioned here that Lifshitz (1948) and Ivanenko and Sokolov (1948) have given calculations leading to lifetimes for the triplet bound state of $8{\cdot}8\times10^{-8}$ sec and $6{\cdot}4\times10^{-7}$ sec respectively.

The decay time in oxygen of 8×10^{-8} sec at 2·4 atmospheres is not very different from that of triplet positronium. The lifetime of γ-emission from triplet positronium would be decreased if transitions occurred to the singlet state (which latter annihilate in 10^{-10} sec). The slight decrease would imply a small number of such transitions. But the spectrum associated with this period showed no sign of triplet state. It would therefore appear that the 8×10^{-8} sec period does not involve positronium, but that it arises from positrons decaying by collisions.

Acknowledgments

The authors wish to thank Professor P. I. Dee for his interest in the work ; also the Department of Scientific and Industrial Research for a grant to one of us (A.T.G.F.).

References

Bell, R. E., Graham, R. L., and Petch, H. E., 1952, *Can. Journ. Phys.*, **30**, 35.
de Benedetti, S., and Siegel, R., 1952, *Phys. Rev.*, **85**, 371.
Deutsch, M., 1951, *Phys. Rev.*, **83**, 866.
Deutsch, M., and Dulit, E., 1951, *Phys. Rev.*, **84**, 601.
Dirac, P. A. M., 1930, *Proc. Camb. Phil. Soc.*, **26**, 361.
Heitler, W., 1944, *The Quantum Theory of Radiation* (Oxford : University Press), p. 216.
Ivanenko, D., and Sokolov, A., 1948, *Doklady Akad. Nauk, S.S.S.R.*, **61**, 51.
Kubitschek, H. E., 1950, *Phys. Rev.*, **79**, 23.
Lifshitz, E. M., 1948, *Doklady Akad. Nauk, S.S.S.R.*, **60**, 211.
Ore, A., and Powell, J. L., 1949, *Phys. Rev.*, **75**, 1696.
Wheeler, J. A., 1946, *Ann. New York Acad. Sci.*, **48**, 219.

CVI. *Isotopic Spin Selection Rules—II : Electric Dipole Transitions without Change of Isotopic Spin in Nuclei of* $T_z = +1$

By D. H. Wilkinson
Cavendish Laboratory, Cambridge*

[Received May 25, 1953]

Abstract

Electric dipole transitions between states of the same isotopic spin in self-conjugate nuclei show very much reduced radiative widths as predicted by the isotopic spin selection rule ($\Delta T = \pm 1$ when $T_z = 0$). In order to show that this inhibition is due to the isotopic spin rule it must be demonstrated that such transitions are not discouraged in nuclei of $T_z \neq 0$. This is done by comparing the cross section of radiative capture of thermal neutrons in ^{7}Li, ^{9}Be and ^{13}C with theoretical values derived with the aid of 'uninhibited' values of the electric dipole radiative widths.

§ 1. Introduction

It has recently been pointed out (Radicati 1952, Gell-Mann and Telegdi 1953) that, if specifically-nuclear interactions are charge-independent, electric dipole transitions are strongly discouraged between states of the same isotopic spin in self-conjugate nuclei ($T_z = 0$). Owing to the Coulomb perturbation isotopic spin states are not quite pure and Radicati (1953) has shown that the factor of discouragement afforded in practice by the isotopic spin selection rule is probably of order 10^2 to 10^4 for transitions not involving states of too-high excitation. That E1 transitions without change of isotopic spin in nuclei of $T_z = 0$ are indeed strongly discouraged has been demonstrated in several instances (Wilkinson and Jones 1953, Jones and Wilkinson 1953, and forthcoming papers in this present series). It is natural to ascribe this discouragement to the operation of the isotopic spin selection rule. Before we may do this, however, it is necessary to show that such E1 transitions, forbidden in self-conjugate nuclei, are not equally discouraged in neighbouring isobars that are not self-conjugate ($T_z = \pm 1$). Ideally we should compare the radiative width of, say, a $T = 1$ to $T = 1$ transition in a nucleus of $T_z = 0$ with the widths of the same transitions in the neighbouring isobars. This has so far proved impossible, but we may derive the same assurance by demonstrating that any E1 transitions without change of isotopic spin are allowed in nuclei of $T_z = \pm 1$.

*Communicated by the Author.

§ 2. The Radiative Capture of Thermal Neutrons

We desire to determine the strength of E1 transitions without change of isotopic spin in nuclei of $T_z=\pm 1$. This is not possible for nuclei of $T_z=-1$ (taking the z-component of the neutron's isotopic spin as positive) since no suitable nucleus of $T_z=-\frac{1}{2}$ exists to serve as target for a (pγ) reaction. For nuclei of $T_z=+1$ the situation is better; we must investigate suitable (nγ) reactions using target nuclei of $T_z=+\frac{1}{2}$. Unfortunately no suitable (nγ) resonances of established characteristics are known and it may seem that this approach may not be followed for the present. The situation is saved, however, by our only needing the assurance that the transitions are not discouraged by a large factor; we may use somewhat approximate arguments relating to the radiative capture of thermal neutrons.

When a thermal neutron approaches a target nucleus of spin I and odd mass we are assured that the parity of the resultant state is that of the target nucleus and that its spin is $I\pm\frac{1}{2}$. Thus, beyond the alpha-particle and until we have filled the $1p_{1/2}$ neutron shell (at ^{14}C, ^{15}N or ^{16}O) the addition of a thermal neutron to a nucleus of $T_z=+\frac{1}{2}$ will give a state of $T_z=+1$, $T=1$ of parity opposite to that of the $T=1$ ground state ultimately reached. There are four target nuclei, ^{7}Li, ^{9}Be, ^{11}B and ^{13}C, all of odd parity with respective spins 3/2, 3/2, 3/2, and 1/2; the ground states of the final nuclei ^{8}Li, ^{10}Be, ^{12}B and ^{14}C all have even parity (presumptive for ^{12}B) and respective spins 2, 0, unknown and 0. E1 radiation to the ground state is therefore possible following thermal neutron capture in ^{7}Li, ^{9}Be, ^{13}C and possibly ^{11}B. In the case of ^{7}Li it is certain that the capture radiation is E1—§ 7 (*a*)—although the transition has not been directly observed; in the case of ^{9}Be(nγ)^{10}Be the E1 transition to the ground state is observed as is another E1 transition to the first excited state (2+); the reaction ^{11}B(nγ)^{12}B is not known; thermal neutron capture in ^{13}C is known but the spectrum has not been measured.

It appears that four E1 transitions are known linking states of $T=1$ in nuclei of $T_z=+1$ and we must ask if their strengths are what we should expect of uninhibited E1 transitions.

Although we may not properly apply the Breit–Wigner formula when remote from the resonance levels concerned the error introduced by so doing is not likely to be of great importance.* We write for the cross section $\sigma_{\gamma x}$ for radiative capture leading to level x of the residual nucleus and due to a level at E_r of spin J:

$$\sigma_{\gamma x}=\frac{\lambda^2}{4\pi}\frac{(2J+1)}{(2I+1)}\frac{\Gamma_n\Gamma_{\gamma x}}{(E_{\gamma 0}-E_r)^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

* The whole treatment of this paper is in the spirit of conventional 'resonances'. This is done, as will be seen, in order to make a prediction of the thermal neutron capture cross section that is free of guessed constants. We are, however, in a region where s-wave nucleon interactions behave very much like single-particle interactions and it may well be more realistic to treat the radiative capture as a one-stage than as a two-stage process. This will receive further attention in § 7.

$E_{\gamma 0}$ is the excitation produced on adding a thermal neutron to the target nucleus of spin I.

We will write Γ_n in terms of the reduced width $\gamma_n{}^2$ of the level at E_r

$$\Gamma_n = 4\pi\gamma_n{}^2/\lambda. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

If now we can decide on the location of the levels dominating thermal capture, their reduced widths, and the expected ' uninhibited ' value of $\Gamma_{\gamma x}$ we may compare the theoretical value of $\sigma_{\gamma x}$ with the experimental value. If the two values are in rough agreement we have established that E1 transitions without change of isotopic spin are not discouraged in nuclei of $T_z = +1$; if such transitions are as strongly inhibited as in nuclei of $T_z = 0$ the experimental value of $\sigma_{\gamma x}$ will be 10^2 to 10^4 times smaller than the theoretical value.

§ 3. The Levels Involved

We must examine the nuclei of interest from the point of view of levels of spin $I \pm \frac{1}{2}$ and odd parity such as may participate in the thermal interaction and of levels that may combine with these by E1 radiation :

(*a*) *Lithium* 8. $E_{\gamma 0} = 2{\cdot}04$ MeV. Little is known of the level structure of this nucleus (see Ajzenberg and Lauritsen (1952) for all references to level schemes) but we need not concern ourselves with it in detail—§ 7 (*a*).

(*b*) *Beryllium* 10. $E_{\gamma 0} = 6{\cdot}81$ MeV. The ground state is of even parity, presumably (0+) ; the first excited state is also of even parity and almost certainly (2+). We now seek states of (1−) and (2−). At 6·2 MeV there is reported a state of undetermined characteristics that may well correspond to the broad (2−) state of ^{10}B at 7·48 MeV. This latter state is in all probability of $T=1$ because it enjoys a large E1 width to the (3+) ground state of ^{10}B. The first $T=1$ state of ^{10}B is at 1·74 MeV ; the difference between 7·48−1·74=5·74 MeV and 6·2 MeV may be accounted for by the Thomas shift (Thomas 1952). Although the evidence of thermal neutron scattering—§ 4 (*c*)—strongly suggests the existence of a (1−) state in ^{10}Be close by this presumed (2−) state there is no known state in ^{10}B that may be immediately associated with it. A possible (1−) state is to be found at 6·89 MeV but its strong decay by alpha-particle and deuteron emission suggests that it is of $T=0$. We are, however, in a region of broad and overlapping resonances in ^{10}B and the (1−) $T=1$ state that we seek may well remain hidden. As the energy of these presumed states in ^{10}Be is rather close to $E_{\gamma 0}$ it is probable that the thermal capture is due to them. No other states are known below $E_{\gamma 0}$.

(*c*) *Boron* 12. $E_{\gamma 0} = 3{\cdot}36$ MeV. The reaction ^{11}B(nγ)^{12}B is not known so we do not concern ourselves with any details of this nucleus—§ 7 (*c*).

(*d*) *Carbon* 14. $E_{\gamma 0} = 8{\cdot}17$ MeV. The ground state is (0+). There is some suggestion that there may be a further (0+) level at about

4·1 MEV ; we will not dwell on this possibility since the spectrum of the capture radiation is not known and admission of the level would only increase the theoretical value of σ_γ by about 10%. We now seek states of (0−) and (1−). A level at 6·10 MEV is quite probably (1−) ; it may well correspond with the 8·06 MEV (1−)* level in ^{14}N (of $T=1$ on account of its strong E1 transition to the (1+) ground state). The first $T=1$ level in ^{14}N is at 2·31 MEV ; the difference between $8{\cdot}06-2{\cdot}31=5{\cdot}75$ MEV and 6·10 MEV is well accounted for by the Thomas shift (Thomas 1952, Thomas and Lauritsen 1952). A level at 8·70 MEV in ^{14}N is (0−) or (1−) and of $T=1$—again because of a strong E1 transition to the ground state ; we may consider the possibility of (1−) as unlikely in view of the large reduced widths displayed by the nearby (1−) level ($2{\cdot}1\times10^{-13}$ MEV-CM (Teichmann and Wigner 1952)) and that now in question ($1{\cdot}5\times10^{-13}$ MEV-CM)—the relation of § 4 (*a*) suggests an associated spacing of about 6 MEV for levels of the same spin and parity. It is therefore possible that this 8·70 MEV level in ^{14}N is (0−) and we then expect a (0−) level to appear in ^{14}C shortly above the 6·10 MEV (1−) level. Such a (1−)(0−) doublet is again suggested by the scattering of thermal neutrons in ^{13}C—§ 4 (*c*). As the (1−) level in ^{14}C is rather closer to $E_{\gamma 0}$ than half the expected spacing between (1−) levels we may ascribe thermal neutron capture to it. No other states are known below $E_{\gamma 0}$.

§ 4. THE REDUCED WIDTHS

The reduced widths γ_n^2 are difficult to estimate with accuracy but we have four independent approaches to the problem :

(*a*) We may use the general relation† (see, for example, Blatt and Weisskopf (1952))

$$\gamma_n^2 \sim \frac{D}{\pi K} ,$$

where D is the spacing between levels of the same spin and parity and K is the wave-number of a neutron inside the nucleus ($\sim10^{13}$ cm^{-1}). D is not very well defined for the nuclei of interest to us now ; it appears to lie between 1 and 10 MEV. Let us say $D\sim5$ MEV and so obtain our first estimate : $\gamma_{n1}^2\sim2\times10^{-13}$ MEV-CM.

(*b*) We may express the reduced width in terms of the Wigner limit (Teichmann and Wigner 1952)

$$\gamma_n^2 \leqslant \frac{3\hbar^2}{2Ma}$$

where M is the reduced mass and a is the channel radius. Now the reduced widths for s-particle interactions in the light nuclei ($A<16$)

* I am grateful to Dr. R. G. Thomas for the information that E. Milne of Cal. Tech. has shown this ^{14}N level to have spin 1 by elastic proton scattering in ^{13}C.

† This is strictly a statistical relation and cannot be expected to give very accurate results in the region of interest here where the independent particle model has fair validity ; but its estimate will not be grossly wrong.

are generally large and range from 3% to as much as 50% of the above limit. A very rough idea of γ_n^2 may then be obtained by taking an average value of about 20% of the limit. This suggests $\gamma_{n2}^2 \sim 3 \times 10^{-13}$ MEV-CM.

(*c*) The scattering cross section σ_s for thermal neutrons is compounded of potential or ' hard sphere ' scattering by the nucleus of radius R and the effects of resonances. If the effects of resonances are small $\sigma_s \sim 4\pi R^2$ and a scattering cross section substantially in excess of this figure may enable us to draw some conclusions about the character of the levels that make it so. In general, since we are concerned with target nuclei of odd mass, we shall have two sets of levels concerned in the scattering—of spins $I \pm \frac{1}{2}$. In general the scattering lengths for these two values of the channel spin will be different and so the coherent and total scattering cross sections will differ. The problem of ^{8}Li is complicated but we do not have to concern ourselves with it here—§ 7 (*a*). The coherent cross-section of ^{11}B is not known but again we need not concern ourselves —§ 7 (*c*). The coherent scattering cross section of ^{9}Be is $7{\cdot}54 \pm 0{\cdot}07$ barns with a positive scattering length (see Hughes *et al.* (1952) for most values of cross sections used here) ; the total (bound) cross section is the same—the spin-dependent scattering is less than 0·03 barns (Palevsky and Smith 1952). The coherent scattering cross section of ^{13}C is $4{\cdot}5 \pm 0{\cdot}6$ barns with a positive scattering length ; the total (bound) cross section is $5{\cdot}5 \pm 1{\cdot}0$ barns.

It therefore appears that, for both ^{9}Be and ^{13}C, the scattering lengths are sensibly independent of channel spin ; they are greater than the nuclear radius and, since they are positive, we must invoke the play of bound levels of both possible spins in each nucleus. As we are ignorant of the locations of one of the two levels concerned in each nucleus—§§ 3(*b*), 3(*d*)—we must lump them together, give them a common reduced width and say :

$$\sigma_s = 4\pi \left\{ R + \frac{\gamma_n^2}{E_{\gamma 0} - E_r} \right\}^2 .$$

Our third estimate of the reduced width is then

$$\gamma_{n3}^2 = \left\{ \left(\frac{\sigma_s}{4\pi} \right)^{1/2} - R \right\} (E_{\gamma 0} - E_r) \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

In the case of ^{10}Be we may imagine that, since the triplet and quintuplet scattering lengths are so closely the same, the level that we have noted —§ 3 (*b*)—at 6·2 MEV as a possible (2—) state is, in fact, a doublet (1—)(2—) ; * we attribute all the excess scattering to this doublet and say $E_{\gamma 0} - E_r = 0{\cdot}6$ MEV. In the case of ^{14}C we have noted—§ 3 (*d*)—the probable (1—) level at 6·10 MEV with the good probability of a (0—)

---

* This possibility has been pointed out by Dr. R. G. Thomas in a private communication cited by Ajzenberg and Lauritsen (1952).

level to be found nearby ; we then set $E_{\gamma 0}-E_r=2{\cdot}1$ Mev (again attributing all excess scattering to two levels only).

The estimate of $\gamma_n{}^2$ based on (3) depends on the chosen value of R which it is now rather more important to know accurately than it was in § 4 (*b*). It appears, from several lines of evidence, that the value of R effective in ^{9}Be is rather greater than that given by the conventional expression

$$R=1{\cdot}5\times A^{1/3}\times 10^{-13}\text{ cm}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

In particular, the work of Coon, Graves and Barschall (1952) on the scattering of 14 Mev neutrons shows that a value about 1·25 times as great as indicated by (4) may be appropriate for ^{9}Be ; we therefore use $R=3{\cdot}90\times 10^{-13}$ cm. (It is possible that, for slow neutrons, the effective value of R should be even larger since it is due, in part, to the loosely-bound last $1p_{3/2}$ neutron.) The fast neutron scattering suggests that, for ordinary carbon, the estimate of (4) should be increased by about 10% and we use $R=3{\cdot}90\times 10^{-13}$ cm. (Again ^{13}C may have a rather larger effective radius for thermal neutrons since its last neutron is also rather loosely bound.)

These estimates of $E_{\gamma 0}-E_r$ and R yield, with (3), for ^{10}Be,

$$\gamma_{n3}{}^2=2{\cdot}3\times 10^{-13}\text{ Mev-cm} ;$$

for ^{14}C we have

$$\gamma_{n3}{}^2=5{\cdot}1\times 10^{-13}\text{ Mev-cm}.$$

As we have remarked, the values of R that we have used may be too small and these values of $\gamma_n{}^2$ consequently too large.

(*d*) The reduced width of the corresponding level in the isobar of $T_z=0$ may be known ; we may then take it that that of the level of interest in the isobar of $T_z=+1$ will be twice as great.* We have noted—§ 3 (*b*)—how we may identify the 6·2 Mev level of ^{10}Be with the 7·48 Mev (2—) level of ^{10}B ; the latter has a reduced width of $0{\cdot}50\times 10^{-13}$ Mev-cm (Teichmann and Wigner 1952) so we may s=ay $\gamma_{n4}{}^2 1{\cdot}0\times 10^{-13}$ Mev-cm for ^{10}Be (assuming the reduced widths to be the same for both (1—) and (2—) levels as was suggested by the evidence of thermal neutron scattering—§ 4 (*c*)). The 6·10 level of ^{14}C has been identified with the 8·06 Mev (1—) level of ^{14}N—§ 3 (*d*)—whose reduced width is $2{\cdot}06\times 10^{-13}$ Mev-cm (Teichmann and Wigner 1952) so we may say $\gamma_{n4}{}^2=4{\cdot}1\times 10^{-13}$ Mev-cm for the (1—) level of ^{14}C that is of interest for radiative capture. (We may note, in passing, that the reduced width ($1{\cdot}5\times 10^{-13}$ Mev–cm) of the probable (0—) $T=1$ level at 8·70 Mev in ^{14}N whose partner in ^{14}C we have supposed, with the 6·10 Mev (1—) level in ^{14}C, to be responsible for the thermal neutron scattering by ^{13}C is closely similar to that of the 8·06 Mev level in ^{14}N and, when transferred to ^{14}C supports our giving the two levels a common reduced width in § 4 (*c*).)

---

* I am indebted to Mr. A. M. Lane for the remark that the reduced width in the $T_z=+1$ nucleus is not equal to that in the $T_z=0$ nucleus but rather twice as great.

We now summarize our estimates of the reduced widths (in 10^{-13} MEV-CM) in the two nuclei of interest :

Nucleus	γ_{n1}^2	γ_{n2}^2	γ_{n3}^2	γ_{n4}^2
^{10}Be	2	3	2·3	1·0
^{14}C	2	3	5·1	4·1

We have remarked—§ 4 (*c*)—that γ_{n3}^2 may be too large. Considering the crudity of the first three estimates we may say that these values are remarkably consistent. We shall use γ_{n4}^2 since it appears to be the most accurate.

§ 5. The Radiative Widths

For the radiative widths $\Gamma_{\gamma x}$ we may make use of the estimates of the single particle model (Weisskopf 1951). It appears, from a survey of E1 radiative widths in light nuclei (Wilkinson 1953) that the estimates of this model are surprisingly reliable, if slightly modified, independently of the type of nucleus concerned, and we use the suggestion of that survey, namely

$$(2J+1)\Gamma_{\gamma x} \sim 0{\cdot}022 A^{2/3} E_{\gamma x}^3 \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

where $\Gamma_{\gamma x}$ is expressed in ev and $E_{\gamma x}$ is the energy in MEV of the gamma-ray produced in the transition to level x.

§ 6. The Radiative Capture Cross Section

We now combine (1), (2) and (5) and insert the value of λ appropriate to thermal neutrons to find

$$\sigma_{\gamma x} = 0{\cdot}017 \frac{\gamma_n^2 A^{2/3} E_{\gamma x}^3}{(2I+1)(E_{\gamma 0}-E_r)^2} \text{mb} \quad . \quad . \quad . \quad . \quad . \quad (6)$$

where the energies are expressed in MEV and γ_n^2 in 10^{-13} MEV-CM.

§ 7. Comparison with Experiment

We shall now examine the elements in turn :

(*a*) *Lithium.* $\sigma_\gamma = 33 \pm 5$ mb. The spectrum of the capture radiation is unknown. It has long been recognized that this cross section is anomalously high ; for example, with $E_{\gamma 0} - E_r = 1$ MEV (and there is no reason to suppose it to be less) (6) suggests less than a millibarn. Thomas (1951) has shown that the high cross section is due to the *extra*-nuclear contribution to the E1 transition moment which is not taken account of by (1) ; his calculated cross section is of the order 30 mb and it is clear that this E1 transition suffers no inhibition.

(*b*) *Beryllium.* $\sigma_\gamma = 10 \pm 1$ mb. Bartholomew and Kinsey (1953) have shown that 0·75 of the capture processes lead directly to the ground state ($\sigma_{\gamma 0} = 7{\cdot}5$ mb) while 0·25 of them lead to the (2+) state at 3·37 MEV

($\sigma_{\gamma 3\cdot 37}=2\cdot 5$ mb). (6) suggests $\sigma_{\gamma 0}=17$ mb ; $\sigma_{\gamma 3\cdot 37}=4$ mb (using both initial levels in ^{10}Be for $\sigma_{\gamma 3\cdot 37}$). This agreement is good and it is clear that these E1 transitions have not been discouraged.

(*c*) *Boron.* $\sigma_\gamma<50$ mb. (6) suggests $\sigma_\gamma\sim 1$ mb so that the present limit upon σ_γ cannot be taken as evidence for a diminished E1 matrix element.

(*d*) *Carbon.* $\sigma_\gamma=1\cdot 0\pm 0\cdot 3$ mb (according to Hughes *et al.* (1952) ; Way *et al.* (1950) quote $\sigma_\gamma=100$ mb).* Nothing is known about the spectrum of capture radiation. (6) suggests $\sigma_\gamma=25$ mb. This figure is substantially greater than the more reliable of the above experimental values but perhaps may not be said to exceed it by more than the uncertainties of the calculation. As $E_{\gamma 0}-E_r=2\cdot 1$ Mev is here comparable with our rough estimate of D—§ 4 (*a*)—(in contrast to the supposed situation in beryllium) we must suspect that other ^{14}C levels may be of importance in this capture process ; they may be of opposite phase at $E_{\gamma 0}$ to the 6·1 Mev level and so tend to diminish σ_γ. We must also bear in mind the possibility of a reduction in the cross section due to the extra-nuclear contribution such as is observed in ^{12}C(nγ)^{13}C (Thomas 1952).†

It appears from these four examples in ^{8}Li, ^{10}Be and ^{14}C that E1 transitions without change of isotopic spin are not inhibited in nuclei of $T_z=+1$ and that we may, in consequence, safely attribute the low radiative widths persistently observed for such transitions in nuclei of $T_z=0$ to the operation of the isotopic spin selection rule.

Acknowledgment

I am very much indebted to Mr. A. M. Lane for reading an early draft of this paper and for making the valuable contributions that are indicated in the text.

---

* I am grateful to Dr. Hughes for a communication in which he expresses confidence in the lower value of σ_γ.

† We have already noted—footnote to § 2—that a single-particle model may afford a more accurate approach to the present problem than that of conventional resonances adopted here. This has been stressed by Mr. A. M. Lane who has kindly made some estimates of the present cross sections on the basis of such a model ; they do not differ materially from those derived above and may be taken to support the conclusion that there is no inhibition of the El transitions in these nuclei of T_z $+1$. In particular it emerges from these calculations that the external contribution to the transition moment in ^{14}C may well be enough to diminish σ_γ to the observed low figure.

Lane has also pointed out that the effective position of a bound level when we wish to insert it into a resonance formula such as (1) is not the observed position of the level but rather lower (perhaps by 1 Mev). This is because the actual bound state wave-function must join to an exponentially-decreasing wave-function outside the nucleus and cannot display the horizontal tangent at the nuclear surface that characterizes the position of a virtual resonance level ; it is then easy to see that the effective position of a bound level is always lower in energy than the actual position. This means that the value of $E_{\gamma 0}-E_r$ that should be inserted in the final expression (6) for $\sigma_{\gamma x}$ should be greater than that actually used ; this will lower all the theoretical values of $\sigma_{\gamma x}$ somewhat and so improve the agreement between theory and experiment. Unfortunately we are not able to make this correction quantitatively at this time.

References

Ajzenberg, F., and Lauritsen, T., 1952, *Rev. Mod. Phys.* **24,** 321.
Bartholomew, G. A., and Kinsey, B. B., 1953, *Canad. J. Phys.*, **31,** 49.
Blatt, J. M., and Weisskopf, V. F., 1952, *Theoretical Nuclear Physics* (Wiley).
Coon, J. H., Graves, E. R., and Barschall, H. H., 1952, *Phys. Rev.*, **88,** 562.
Gell-Mann, M., and Telegdi, V. L., 1953, *Phys. Rev.* (in press).
Hughes, D. J., *et al.*, 1952, AECU-2040.
Jones, G. A., and Wilkinson, D. H., 1953, *Phys. Rev.*, **90,** 722.
Palevsky, H., and Smith, R. R., 1952, *Phys. Rev.*, **86,** 604.
Radicati, L. A., 1952, *Phys. Rev.*, **87,** 521 ; 1953, *Proc. Phys. Soc.* A, **66,** 139.
Teichmann, T., and Wigner, E. P., 1952, *Phys. Rev.*, **87,** 123.
Thomas, R. G., 1951, *Phys. Rev.*, **84,** 1061 ; 1952, *Ibid.*, **88,** 1109.
Thomas, R. G., and Lauritsen, T., 1952, *Phys. Rev.*, **88,** 969.
Way, K., *et al.*, 1950, *Nat. Bur. Stds. Circular* 499.
Weisskopf, V. F., 1951, *Phys. Rev.*, **83,** 1073.
Wilkinson, D. H., 1953, *Phil. Mag.*, **44,** 450.
Wilkinson, D. H., and Jones, G. A., 1953, *Phil. Mag.*, **44,** 542.

CVII. *Energy Dependence of Neutron Total Cross Sections*

By F. Mandl and T. H. R. Skyrme
Atomic Energy Research Establishment, Harwell*

[Received June 1, 1953]

Abstract

The observed energy dependence of neutron total cross sections in the energy range 60 to 153 mev is analysed in terms of single particle scattering by a complex square well potential, using a first approximation to a variational method. The observed peaks and dips in the cross section can be explained with a reasonable choice of potential.

§ 1. Introduction

Accurate measurements of neutron total cross sections between 30 and 153 mev have recently been carried out (Taylor and Wood 1953). For the heavy elements they show a marked dip and corresponding peak. For lead the dip occurs at about 60 mev, the peak at about 90 mev. For lighter elements, such as copper or cadmium, these are displaced to lower energies. For light elements, such as aluminium, they are altogether absent.

In this paper, we analyse these cross sections in terms of a square well model of the nucleus with a complex potential, the real part describing the elastic or diffraction scattering and the imaginary part describing the inelastic or absorption processes. As the well radius must be essentially the nuclear radius we have two parameters, which may depend on the bombarding energy and the target element, to fit the experimental cross sections. As the nucleus almost certainly has not got a sharp boundary a square well potential will exaggerate surface effects, such as reflection.

In § 2 the solution of this scattering problem by the geometrical optical model of Fernbach *et al.* (1949) is considered. It is found that this approximation does not seem able to explain the experimental results. Qualitatively, however, the model is capable of giving a simple physical reason for the presence of peaks and dips at about the energies observed (Lawson 1953). This is not surprising as geometrical optics might only be adequate at energies of 200 mev or above. In § 3 a variational method of solving the scattering problem is developed and in § 4 it is applied to this problem. It is found that the experimental cross sections can be explained moderately well with a reasonable choice of the complex potential.

* Communicated by the Authors.

§ 2. The Geometrical Optical Method

We describe the field which the incident neutron experiences by a complex square well potential

$$\left.\begin{aligned} V(r)&=0 \qquad & r>R \\ V(r)&=V+iV' \qquad & r<R \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad (2.1)$$

where $R=1{\cdot}37\times10^{-13}A^{1/3}$ is the nuclear radius. This potential corresponds to a complex refractive index

$$\frac{k'}{k}=1+\frac{1}{k}\left(k_1+\frac{iK}{2}\right), \quad . \quad . \quad . \quad . \quad . \quad . \quad (2.2)$$

where k and k' are the wave numbers of the neutron, whose incident energy is E, outside and inside the nucleus, i.e.

$$\left.\begin{aligned} k&=(2mE/\hbar^2)^{1/2} \qquad & r>R \\ k'&=(2m[E-V-iV']/\hbar^2)^{1/2}. \qquad & r<R \end{aligned}\right\} \quad . \quad . \quad (2.3)$$

One can solve this scattering problem using the W.K.B. approximation, i.e. considering the limiting case of geometrical optics. Fernbach *et al.* (1949) have given a particularly simple form of this approximation in which all reflection and refraction at the nuclear surface are neglected. The error introduced in this way is probably not serious as the nuclear surface is not a sharp boundary. On the other hand the geometrical optical approximation is only valid at high energies; one would not expect it to be adequate below about 200 Mev, particularly for light elements.

We have tried to fit the theoretical expressions for the cross sections of Fernbach *et al.* to the experimental data of Taylor and Wood for lead and aluminium in the energy range 60 Mev to 153 Mev for suitable values of K and k_1.

For the absorption length or the mean free path $1/K$ it is most reasonable to take a slightly modified form of the original expression of Fernbach *et al.*, which relates K to the free nucleon total cross sections σ_{np} and σ_{nn} (taken equal to σ_{pp}),

$$K=\frac{3}{4\pi R^3}[Z\alpha_{np}\sigma_{np}+N\alpha_{nn}\sigma_{nn}], \quad . \quad . \quad . \quad . \quad . \quad (2.4)$$

where α_{np} and α_{nn} are numbers smaller than unity which take account of the Pauli exclusion principle (Goldberger 1948). In table 1 we give some typical values of K for lead and aluminium.

Existing data on the absorption cross sections of neutrons (De Juren and Knable 1950, De Juren 1950) are admittedly not very satisfactory so that it is difficult to check eqn. (2.4). However, nuclear absorption cross sections for protons above 130 Mev agree quite well with those obtained from (2.4) in conjunction with the geometrical optical model (Kirschbaum 1952, Lees *et al.* 1953, Cassels and Lawson, private communication, to be published).

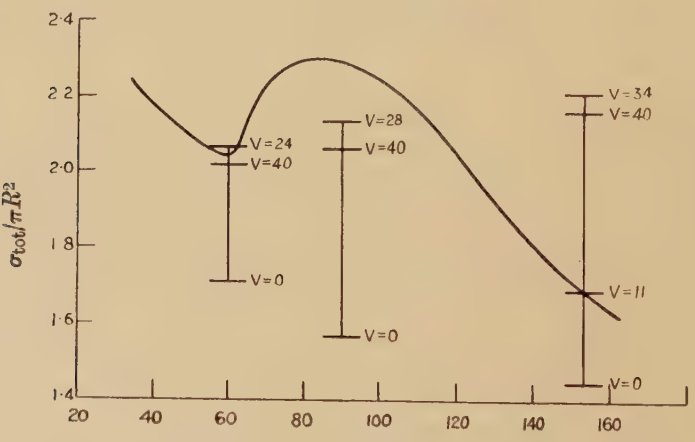
In figs. 1 and 2 we have plotted the experimental total neutron cross sections of Taylor and Wood, as functions of energy, for lead and aluminium respectively. We have also marked the cross sections as deduced from the optical model at neutron bombarding energies of 60, 90 and 153 MeV for different nuclear well depths V. The maximum

Table 1

Neutron energy E MeV	$K.10^{-12}$ cm^{-1}	
	Pb	Al
60	4·3	4·5
90	3·2	3·3
153	2·6	2·7

cross section shown at each energy represents the maximum possible cross section on the geometrical optical model with the above values of K, and varying V will not increase this maximum. Studying figs. 1 and 2 we note that a Fermi model of the nuclear potential, as originally suggested by Fernbach *et al.*, is not satisfactory. We cannot think of

Fig. 1



Energy variation of neutron total cross sections for lead.

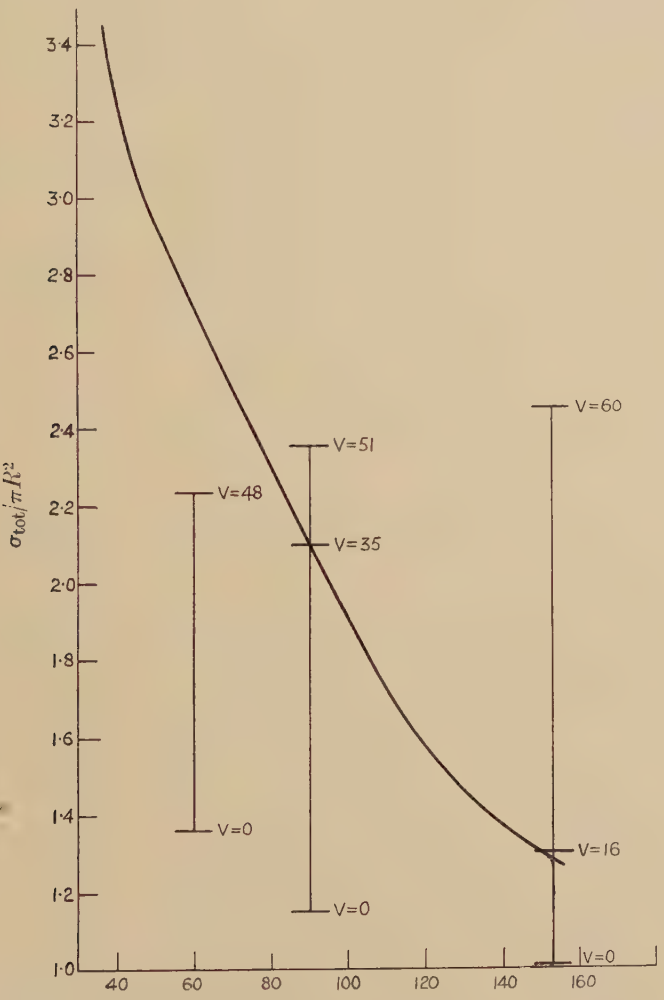
Curve : Experimental values of Taylor *et al.*

Ordinates marked : Theoretical values calculated by the geometrical optical method for a well depth of V MeV.

the nucleus as having a constant well depth, calculated on the model of an electron gas, with a value of about 20 to 30 MeV. Rather it seems as though at higher energies the well depth decreases to about 10–15 MeV. Such deviations from the Fermi model seem reasonable. In our analysis, V represents a sort of mean potential which the incident particle sees

and it is very likely that this is energy dependent. In particular, for spin dependent nucleon–nucleon interactions one might expect the potential to decrease at high energies, as the time which the projectile neutron spends inside the nucleus is too short for any alignment of spins

Fig. 2



Neutron energy in Mev.

Energy variation of neutron total cross sections for aluminium.

Curve : Experimental values of Taylor *et al.*

Ordinates marked : Theoretical values calculated by the geometrical optical method for a well depth of V Mev.

to occur : the incident neutron only sees a resulting mean spin of the nucleus which is small. Secondly we must analyse the fact that the experimental cross sections cannot be reproduced theoretically with our assumptions. For lead we cannot obtain a sufficiently large peak at

90 MEV, for aluminium we cannot obtain a sufficiently large cross section at 60 MEV. This disagreement could be due to the values of K we have chosen or the inadequacy of the optical method. We have carried out some calculations which show that it is not possible to obtain essentially better agreement by changing K. On the other hand, as stated above, we might expect the optical method to fail at the energies considered, and that we should use a wave treatment for this scattering problem.

§ 3. A Variational Method for Calculating Total Cross Sections

Our scattering problem is represented by the following integral equation

$$\psi(\mathbf{r})=\psi_0(\mathbf{r})-\int K(\mathbf{r}-\mathbf{r}')U(\mathbf{r}')\psi(\mathbf{r}')\,d\mathbf{r}', \quad . \quad . \quad . \quad . \quad (3.1)$$

where $\psi_0(\mathbf{r})=\exp(ikz)$ represents the plane incident wave, $\{-U(\mathbf{r})\}$ is the scattering potential in units of $\hbar^2/2m$ and $K(\mathbf{r}-\mathbf{r}')$ is the Green's function of the wave equation,

$$K(\mathbf{r}-\mathbf{r}')=-\frac{1}{4\pi}\frac{\exp\{ik|\mathbf{r}-\mathbf{r}'|\}}{|\mathbf{r}-\mathbf{r}'|}.$$

One can easily show that for the wave function $\psi(\mathbf{r})$ satisfying (3.1) the following functional

$$F[\psi]=\frac{[\int\psi(\mathbf{r})\,U(\mathbf{r})\,\psi_0(-\mathbf{r})\,d\mathbf{r}]^2}{\{\int\psi(\mathbf{r})U(\mathbf{r})\psi(-\mathbf{r})d\mathbf{r}+\int\int\psi(\mathbf{r})U(\mathbf{r})K(-\mathbf{r}-\mathbf{r}')U(\mathbf{r}')\psi(\mathbf{r}')d\mathbf{r}'\,d\mathbf{r}\}} \quad . \quad . \quad . \quad . \quad . \quad . \quad (3.2)$$

assumes a stationary value, $\delta F=0$, provided $U(-\mathbf{r})=U(\mathbf{r})$, which is the case for our potential, eqn. (2.1). For such a stationary solution $F[\psi]$ reduces to $F_0[\psi]$ given by

$$F_0[\psi]=\int\psi(\mathbf{r})U(\mathbf{r})\psi_0(-\mathbf{r})\,d\mathbf{r}, \quad . \quad . \quad . \quad . \quad . \quad . \quad (3.3)$$

and the total cross section of the scattering process is given by the well-known formula

$$k\sigma_{\text{tot}}=ImF_0[\psi]. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3.4)$$

§ 4. Application of the Variational Method to the Scattering Problem

To solve the scattering problem of § 2 by the variational method using trial wave functions containing unknown parameters, requires very heavy numerical analysis. We have therefore not carried out a variational calculation but approximated the cross section (3.4) by

$$k\sigma_{\text{tot}}=ImF[\psi], \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4.1)$$

where for ψ we take

$$\psi(\mathbf{r})=\exp(ik'z) \qquad (r<R), \quad . \quad . \quad . \quad . \quad . \quad . \quad (4.2)$$

as suggested by the optical model. (The wave function for $r>R$ does not enter $F[\psi]$ since U vanishes if $r>R$.) If we take for the wave function

in (4.1) the unperturbed wave $\psi=\psi_0$, (4.1) gives the cross section to at least the accuracy of second Born approximation. Using (4.2) instead of ψ_0 increases the accuracy further. At high energies, where first Born approximation is valid, our method reduces to it ; at low energies it should be considerably better.

To check the variational method we have calculated the total cross section of aluminium for 90 MEV neutrons. In table 2 we compare this value with those obtained by the optical method and by an exact phase calculation of Pasternack and Snyder (1950) using the same data ($K/k=0{\cdot}1$, $k_1/k=0{\cdot}15$, $kR=8{\cdot}56$). We also give the elastic and inelastic cross sections derived by the two latter methods. We see from table 2 that the variational calculation gives a result about 10% too large, the optical method one about 10% too small. The optical method thus

Table 2

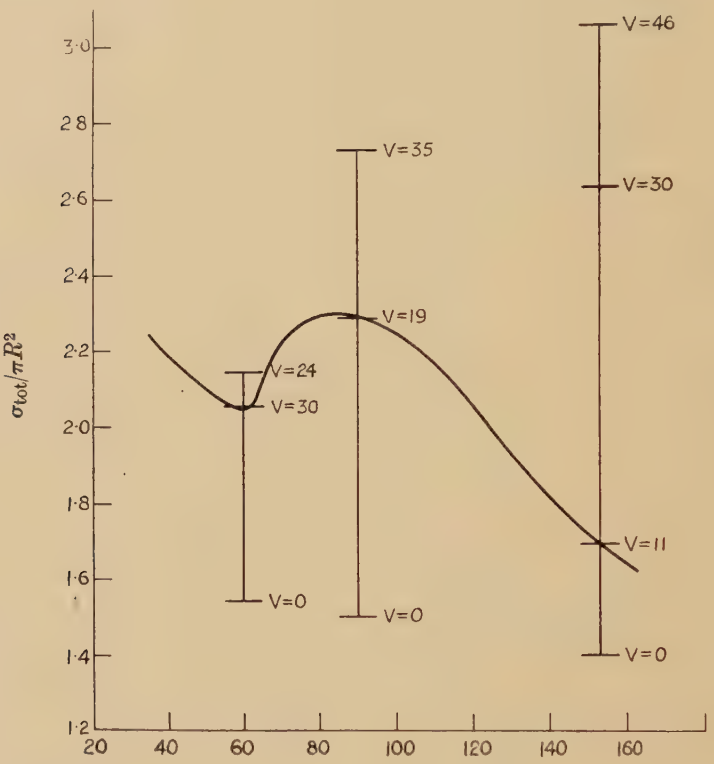
Cross section in barns	Model		
	optical	Pasternack and Snyder	variational
absorption	0·36	0·45	—
elastic	0·76	0·83	—
total	1·12	1·28	1·40

appears better than one might have expected at this energy. As regards the absorption cross section the optical method gives too small a value. This was to be expected as internal reflections are neglected in this treatment and the neutron's path length in the nucleus is underestimated. It is difficult to say anything about the elastic cross section as this depends critically on the phase shifts. It is reasonable that the geometrical optical model gives a better value for the total cross section than for the separate elastic and inelastic cross sections. Unlike the latter, the total cross section depends only on the mean amplitude of the attenuated beam so that interference phenomena, which are neglected in this treatment, are unimportant.

Using eqns. (4.1) and (4.2) we have analysed the experimental cross sections in the same way as in § 2. In figs. 3 and 4 we again give the experimental cross sections for lead and aluminium and the cross sections calculated on the variational method at 60, 90 and 153 MEV for different nuclear well depths. For the absorption length $1/K$ we took the conventional values given by eqn. (2.4). It follows from the figures that good agreement is obtained by the variational method if one takes the nuclear potential V to be a steadily decreasing function of energy, roughly independent of the element, and varying from approximately 30 MEV at neutron energies of 60 MEV to approximately 11 MEV at neutron energies of 153 MEV. Such an energy dependence is not unreasonable, as discussed

in § 2. The value of 11 MeV for the well depth at high energies is only approximate, as the phase retardation of the neutron beam by the nucleus is small at these energies so that the calculation of V is not very reliable. The same criticism applies to the well depth obtained for the optical model in § 2.

Fig. 3



Neutron energy in MeV.

Energy variation of neutron total cross sections for lead.

Curve : Experimental values of Taylor *et al.*.

Ordinates marked : Theoretical values calculated by the variational method for a well depth of V MeV.

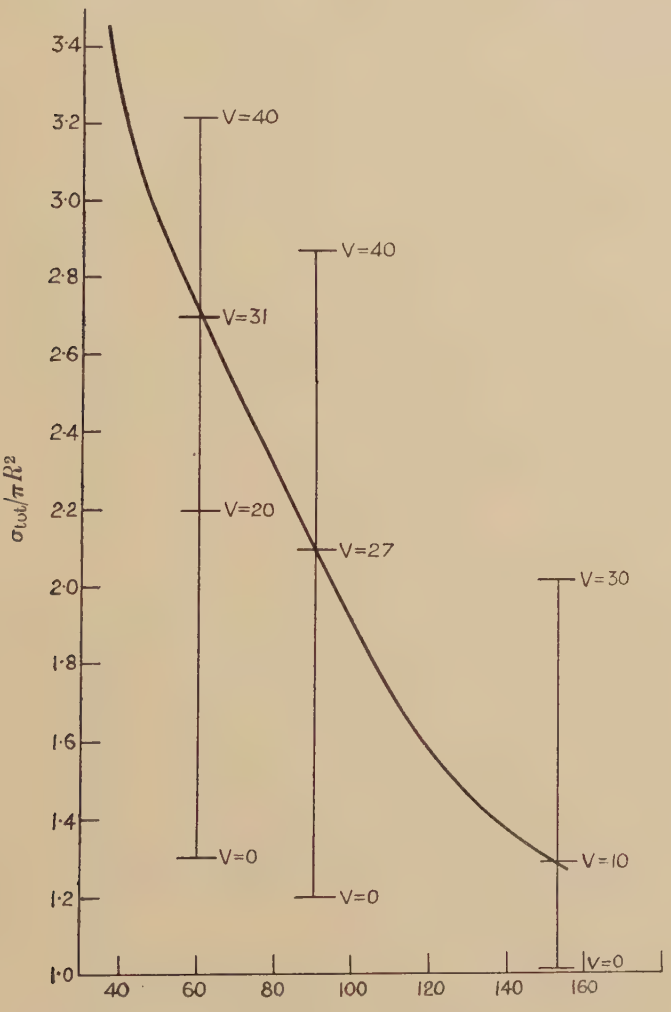
Above we obtained values for the nuclear potential V, i.e. the refractive index $n=1+k_1/k$, which fit the experimental cross sections. It would represent a useful independent check of the theory if one could deduce the refractive index n from the nucleon–nucleon scattering data. Jastrow (1951) has stated a relation for the refractive index n. In a slightly modified form it is

$$n^2-1=\frac{6}{k^2R^3}\{N[\tfrac{3}{4}f^t_{nn}+\tfrac{1}{4}f^s_{nn}]+Z[\tfrac{3}{4}f^t_{np}+\tfrac{1}{4}f^s_{np}]\}, \quad . \quad . \quad . \quad (4.3)$$

where f^t_{np} and f^s_{np} are the n–p triplet and singlet forward scattering amplitudes in the centre of mass system, and f^t_{nn} and f^s_{nn} are similar

quantities for n–n scattering (for either particle to be scattered in the forward direction). This relation is easily derived in Born approximation, treating the particles in the nucleus as independent fixed scattering centres. Using an impulse approximation we can obtain a more general proof but it is difficult to estimate its range of validity.

Fig. 4



Neutron energy in MeV.

Energy variation of neutron total cross sections for aluminium.

Curve : Experimental values of Taylor *et al.*

Ordinates marked : Theoretical values calculated by the variational method for a well depth of V MeV.

Some authors (Jastrow 1951, Memmert 1952) have suggested using eqn. (4.3) to predict the refractive index n but we doubt whether this is possible. Even if one assumes, as is probably correct, equality of the

n–n and p–p interactions, apart from the Coulomb field, Memmert's suggestion, to use the experimental nucleon–nucleon scattering data, is erroneous as these only give the mean square values of singlet and triplet amplitudes. Even these are not known in the forward direction and for protons can probably not be determined on account of the Coulomb scattering. Nor can the forward scattering amplitudes be deduced with any degree of certainty from the present theoretical knowledge of nucleon–nucleon scattering.

We conclude that

(i) it is difficult to interpret the refractive index describing the scattering in terms of the known two-nucleon data,

(ii) the observed total neutron cross sections agree with those calculated from a variational principle with a reasonable choice of refractive index, which is roughly independent of the target element.

Acknowledgments

We are indebted to Dr. J. M. Cassels for stimulating discussions of topics related to the above problem, to Dr. J. M. Cassels and Mr. B. H. Flowers for helpful criticism of this paper and to Mr. S. Outram for carrying out the computations. This paper is published by permission of the Director, Atomic Energy Research Establishment, Harwell.

References

De Juren, J., 1950, *Phys. Rev.*, **80,** 27.
De Juren, J., and Knable, N., 1950, *Phys. Rev.*, **77,** 606.
Fernbach, S., Serber, R., and Taylor, T. B., 1949, *Phys. Rev.*, **75,** 1352.
Goldberger, M. L., 1948, *Phys. Rev.*, **74,** 1269.
Jastrow, R., 1951, *Phys. Rev.*, **82,** 261.
Kirschbaum, A. J., 1952, *University of California Radiation Laboratory Report*, No. 1967.
Lawson, J. D., 1953, *Phil. Mag.*, **44,** 102.
Lees, C. F., Morrison, G. C., Muirhead, H., and Rosser, W. G. V., 1953, *Phil. Mag.*, **44,** 304.
Memmert, G., 1952, *Z. f. Phys.*, **134,** 42.
Pasternack, S., and Snyder, H. S., 1950, *Phys. Rev.*, **80,** 921.
Taylor, A. E., and Wood, E., 1953, *Phil. Mag.*, **44,** 95.

CVIII. *The Determination of the Mass of Energetic Helium Isotopes Emitted in Nuclear Explosions*

By T. Holtebekk, N. Isachsen and S. O. Sörensen
Fysisk Institutt, University of Oslo*

[Received May 18, 1953]

Summary

The mass spectrum of energetic helium isotopes, ejected during the explosive disintegration of nuclei, is determined, and the ^{3}He/^{4}He-ratio is found to be less than 0·05. The frequency of emission of energetic ^{6}He-nuclei is probably negligible. Measurements are carried out on the ratio of single protons to deuterons at an altitude of 3500 m, and the result is found to be in accordance with previous estimates. The identification of the particles depends on a particular scattering method which is described.

§ 1. Introduction

In experiments with photographic plates exposed to cosmic radiation, several reports have been given of the observations of heavy nuclear ' fragments ' ejected during the explosive disintegration of nuclei. The most striking feature of these phenomena is that the heavy particles are frequently emitted with velocities much greater than that which can be accounted for in terms of the electrostatic repulsion of the residual nucleus. The problem therefore arises as to the physical mechanism through which it is possible to endow these nuclei with kinetic energies greatly in excess of their binding energy without disrupting them in the process. The probability of emission of the particles falls off very rapidly with increasing atomic number (Perkins 1950, Crussard 1952) and the most thorough investigation of the phenomena has been confined to the emission of energetic helium nuclei.

The most characteristic feature of the heavy fragments is the large number of ' knock on ' electrons or delta rays which the particles produce in their passage through the emulsion, and a study of the distribution of these delta rays along the tracks provides a precise method for determining the charge of the fragments. The ionization produced by a particle of charge ze is proportional to z^2, and due to the high value, 4, of the z^2-ratio for doubly and singly charged particles, helium nuclei

* Communicated by the Authors.

can easily be distinguished from tracks of hydrogen isotopes by a simple inspection of the delta ray density.

Within the last few years numerous determinations of the masses of particles with unit charge (mesons, hydrogen isotopes) in photographic emulsions have been made by a study of their small angle Coulomb scattering. (cf. Goldschmidt-Clermont *et al.* (1948), Fowler (1950), Menon and Rochat (1951), Daniel *et al.* (1952)). Apart, however, from the well-known lithium fragment $^{8}_{3}$Li, leading to the characteristic ' hammer tracks ' (Franzinetti and Payne 1948), it is in general difficult to obtain an accurate determination of the mass of heavy nuclear fragments.

The present paper describes an experiment to estimate the relative frequency of various isotopes of helium emitted during the explosive disintegration of nuclei. The investigation is based on a particular scattering method using a cell size varying continuously along the track. In order to obtain a sufficient number of independent readings of scattering angles, only helium tracks longer than 1000μ were used, corresponding to alpha particles with an energy greater than $\sim$55 MeV.

§ 2. The Scattering Method

A charged particle traversing a thin layer of matter suffers a deviation which is the resultant of a large number of individual deflections due to elastic collisions with the nuclei of the atoms composing the medium. This scattering is statistical in nature, and various authors have calculated formulae for the probability of a change of direction, Φ, occurring after traversing a given thickness of matter, from the distribution of the values of the angles of scattering occurring in single collisions. Molière (1947, 1948) has given the following formula for the mean of the absolute value of the projection of Φ on a plane containing the incident direction :

$$\bar{\alpha}=\delta \, . \, B^{1/2}(1+0{\cdot}982B^{-1}-0{\cdot}117B^{-2})=\delta \, . \, L \quad . \quad . \quad . \quad . \quad (1)$$

δ is an angular unit given by

$$\delta=\frac{2e^2zt^{1/2}(\Sigma N_i Z_i^{\,2})^{1/2}}{pv} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

where

ze, p, $v=\beta c$=charge, momentum and velocity of the scattered particle.

N, Z=number of scattering centres, usually taken as the number of atoms per cm^3 and the atomic number of the scattering medium. The subscript i refers to the different atomic species.

t=thickness of the scattering medium.

The parameter B is the greater solution of the equation.

$$B-\ln B=\ln \Omega_b-0{\cdot}1544 \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

where

$$\Omega_b=\pi\delta^2/\bar{\chi}^2. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

Molière has made an exact quantum-mechanical calculation of $\bar{\chi}$ and obtains

$$\ln \bar{\chi}^2 = \frac{\Sigma N_i Z_i^2 \ln \chi_i^2}{\Sigma N_i Z_i^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

with

$$\chi_i^2 = \frac{Z_i^{2/3}\hbar^2}{a^2 p^2}(1{\cdot}13+3{\cdot}76\gamma_i^2) \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

where $a=0{\cdot}4685\times 10^{-8}$ cm is the Thomas–Fermi unit of length, and $\gamma_i = Z_i z/137\beta$.

The range–energy relation of a charged particle traversing a photographic plate is given by

$$E = 0{\cdot}251 R^{0{\cdot}575} z^{1{\cdot}15} M^{0{\cdot}425} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (7)$$

where the energy E is expressed in MeV, the mass M in proton masses and the range R measured in microns. By combining (7) with the above formulae, we obtain after simple, non-relativistic calculations

$$\delta = 3{\cdot}486\times 10^{-2} t^{1/2} R^{-0{\cdot}575} z^{-0{\cdot}15} M^{-0{\cdot}425} \quad . \quad . \quad . \quad . \quad . \quad (8)$$

$$\frac{\Sigma N_i Z_i^2 \ln \chi_i^2}{\Sigma N_i Z_i^2} = -7{\cdot}71 + \ln \frac{R^{0{\cdot}575}}{t M^{0{\cdot}575} z^{0{\cdot}85}},$$

$$+ \frac{\Sigma N_i Z_i^2 \ln (1{\cdot}13+0{\cdot}375 Z_i^2 R^{-0{\cdot}575} M^{0{\cdot}575} z^{0{\cdot}85})}{\Sigma N_i Z_i^2}. \quad . \quad . \quad (9)$$

An unambiguous relation therefore exists for various types of particles between the range R and the mean angle of scattering $\bar{\alpha}$. The charge of a particle can be determined, if its range is of sufficient length in the emulsion, by measuring grain densities or delta ray densities along the track. Thus, by measuring R and $\bar{\alpha}$ the mass M of a particle can be determined in principle. This method is necessarily confined to particles arrested in the emulsion.

The accuracy of determining the statistical quantity $\bar{\alpha}$ increases with the number of independent cells. This is obtained by applying tracks of long range and using small cell sizes. The measured scattering of the particle has to be appreciably larger than the spurious scattering (noise) of the track. This sets a lower limit to the shortest cell size, that can be used. In this investigation the following scattering method has been applied :

We express the scattering of a track by the mean deviation $\bar{D}$ given by

$$\bar{D} = \bar{\alpha}_c t, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (10)$$

where $\bar{\alpha}_c$ is the average change of direction between successive chords. As the mean scattering parameter in Molière's theory represents the angle between the tangents, we have $\bar{\alpha}_c = 0{\cdot}816\bar{\alpha}$. From (1), (8) and (10) we obtain

$$\bar{D}(M, z) = 2{\cdot}86\times 10^{-2} t^{3/2} R^{-0{\cdot}575} z^{-0{\cdot}15} M^{-0{\cdot}425} L. \quad . \quad . \quad (11)$$

From (11) we calculate a cell size

$$t=t(R) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (12)$$

which varies along the track of a given type of particles in such a way that the mean deviation is a constant $\bar{D}=\bar{D}_1$. The minimum value which can be chosen for $\bar{D}_1$ is given by the relation

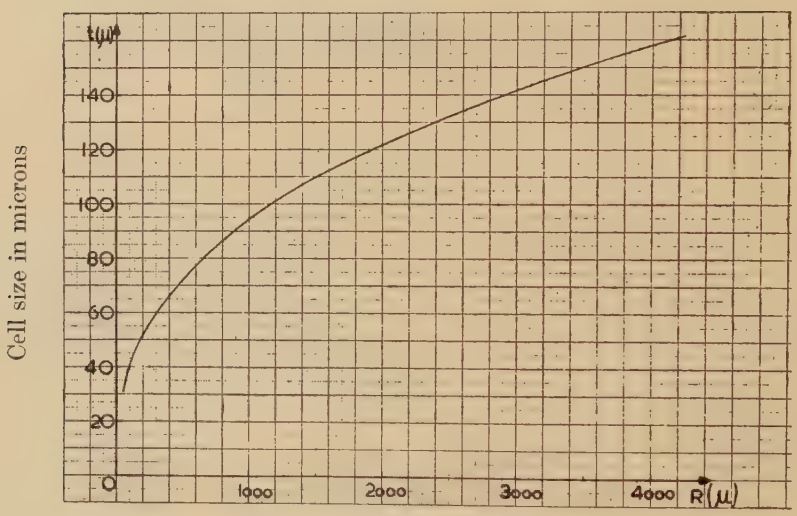
$$\bar{D}_1 \gtrsim 4\bar{D}_{\text{noise}}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (13)$$

The determination of the ' noise ' of the tracks in this investigation gives

$$\bar{D}_{\text{noise}}=(0{\cdot}19\pm 0{\cdot}01)\mu \quad . \quad . \quad . \quad . \quad . \quad . \quad (14)$$

(cf. Appendix). The function (12) is calculated from (11), for an alpha particle ($M=4$, $z=2$) and a mean deviation $\bar{D}(^4_2\text{He})=0{\cdot}80\,\mu$. The result is represented in the curve, fig. 1. The observed mean deviation

Fig. 1



Graph showing the cell size as a function of the residual range. The curve is calculated for an alpha particle and a mean deviation $\bar{D}$ ($^4_2$He)=0·80 μ.

$\bar{D}_{\text{obs}}$ of the track of an alpha particle is then given by

$$\bar{D}_{\text{obs}}(^4_2\text{He})=(\bar{D}^2+\bar{D}^2_{\text{noise}})^{1/2}=0{\cdot}82\,\mu. \quad . \quad . \quad . \quad (15\,a)$$

The observed mean deviation of the other isotopes of helium is calculated from (11), using the same value for the ' noise level '. We find the values

$$\bar{D}_{\text{obs}}(^3_2\text{He})=0{\cdot}92\,\mu, \quad . \quad . \quad . \quad . \quad . \quad . \quad (15\,b)$$

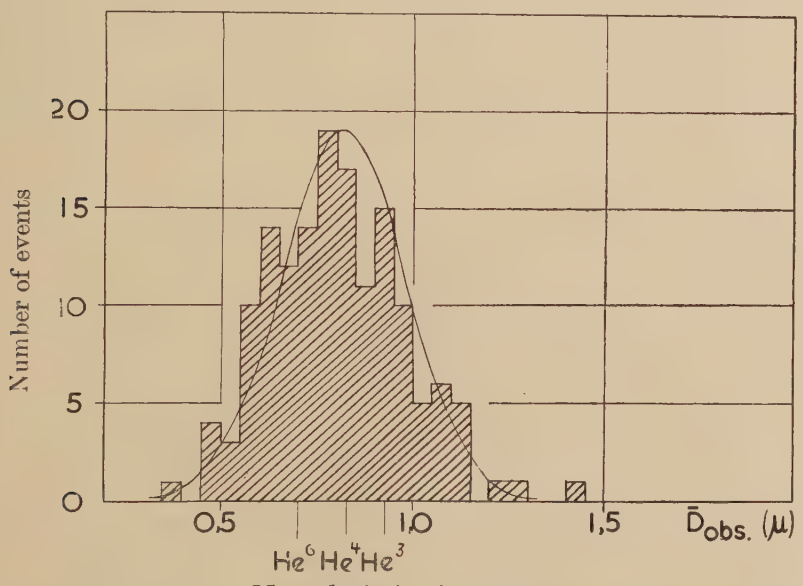
$$\bar{D}_{\text{obs}}(^6_2\text{He})=0{\cdot}69\,\mu. \quad . \quad . \quad . \quad . \quad . \quad (15\,c)$$

$^5_2$He has a short lifetime ($\sim 10^{-20}$ sec) and therefore the question of its observation does not arise.

§ 3. Observations on Helium Tracks

The material used in the present investigation consisted of 13 Ilford G5 plates, 3 in. × 3 in., all 400 μ thick. Free balloons were used for a four hour exposure at an altitude of 90 000 feet. During a systematic scan of the plates 149 helium tracks were found which had been brought to rest in the emulsion and had a residual range greater than 1000 μ. The majority of the particles were ejected from nuclear disintegrations produced by energetic particles in the cosmic radiation. A few tracks entered the emulsion from the surface or the glass and may be a part of the helium component in the primary cosmic ray beam. Further, the particles were estimated to have a charge of $2e$ by visual inspection of the density of delta rays along the tracks. Due to the uncertainty of

Fig. 2



Histogram showing the distribution of $\bar{D}_{obs}$ for 149 helium tracks, when using cell sizes varying according to the curve in fig 1. The full-drawn curve is the normal distribution calculated for $^4_2$He-nuclei.

this method one cannot exclude the possibility of having included in the material a few lithium nuclei of unusually low delta ray density, or particles of unit charge (protons, deuterons and tritons) with very high delta ray density. However, it is reasonable to assume that this effect is negligible.

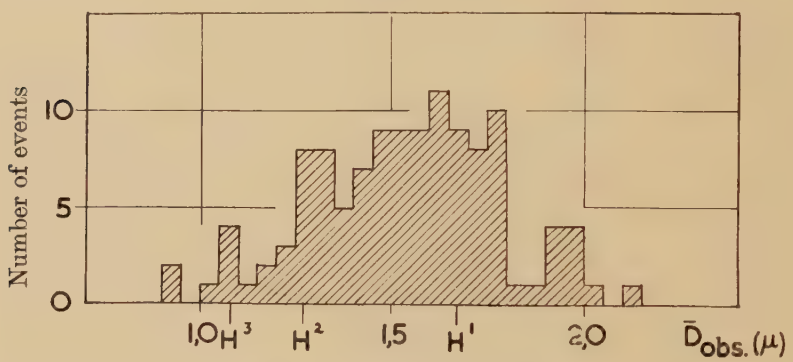
During the observations each ocular had an eye-piece scale, one for measuring the varying cell size and the other for scattering. The ' coordinate method ' described in detail by Fowler (1950) was used in determining the mean deviation $\bar{D}_{obs}$ of the particles. The experimental results are shown in fig. 2. The figure represents a histogram of the

observed values of the mean deviation of the helium tracks, when using cell sizes varying according to the curve in fig. 1. The theoretical values of $\bar{D}_{\text{obs}}$ for the three helium isotopes in question, given by (15 *a*, *b* and *c*), are indicated in fig. 2. The full-drawn curve is the normal distribution calculated for $^4_2$He-nuclei.

§ 4. Observations on Protons and Deuterons

Molière's theory of multiple scattering is still not completely verified by experiment. We therefore have checked the present scattering method with a sample of particles of known relative frequency. As such we have chosen 118 tracks of unit charge, mesons excluded, and minimum range 4000 μ observed in a stack of Ilford G5 plates, exposed on the Jungfraujoch (3500 m). The application of the scattering method on these particles is completely analogous to that outlined for the helium tracks, and the result of the observations is shown in fig. 3. A statistical

Fig. 3



Mean deviation in microns

Histogram showing the distribution of $\bar{D}_{\text{obs}}$ for 118 tracks of particles of unit charge, mesons excluded, observed in plates exposed on the Jungfraujoch.

analysis of the histograms gives the relative frequency of protons and deuterons represented in table 1. The table also contains the result of Franzinetti (1950), on the mass ratio of the two hydrogen isotopes found in magnetic deflection experiments on the Jungfraujoch. The agreement between the results obtained by the two methods is satisfactory.

As an additional check of the present scattering method, the above results on particles of unit charge were compared with those obtained using a method for mass-ratio determinations suggested by Holtebekk (1951) and Dainton and Fowler (1953). The method depends on a study of the delta ray intensity along the track of a particle. This can give both a determination of the charge and an indication of the mass.

The number of delta rays per unit length of the trajectory of a particle is a function of its charge z and velocity v. The various isotopes of an element have different range velocity relations, due to the differences in

mass. Thus, the total number of delta rays for equal residual range will be different for protons and deuterons. This effect will be most pronounced near the end of the tracks where the delta ray density increases strongly with velocity.

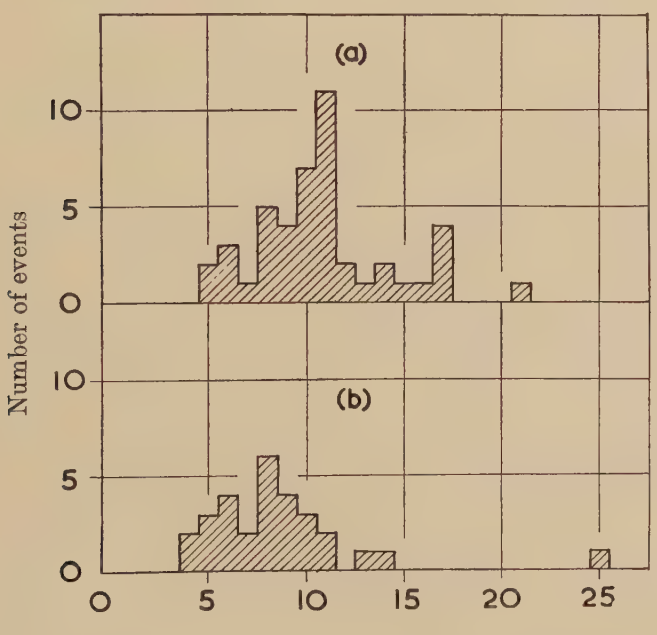
Table 1

Relative Frequency of Slow Protons and Deuterons at 3500 m (Jungfraujoch)

	Scattering measurements	Magnetic deflection experiment	Delta ray measurements
Protons in %	70 ± 8	75 ± 4	72 ± 7
Deuterons in %	30 ± 8	25 ± 4	28 ± 7

From the material represented in fig. 3 we have selected two groups of particles with $\bar{D}_{obs} \geqq 1{\cdot}60\,\mu$ and $\bar{D}_{obs} \leqq 1{\cdot}35\,\mu$ respectively. The total number of delta rays, containing 4 grains or more, in $1400\,\mu$ range

Fig. 4



Number of delta rays.

Histograms showing the distribution in the total number of delta rays, containing 4 grains or more, in $1400\,\mu$ range from the end of the tracks represented in fig 3.

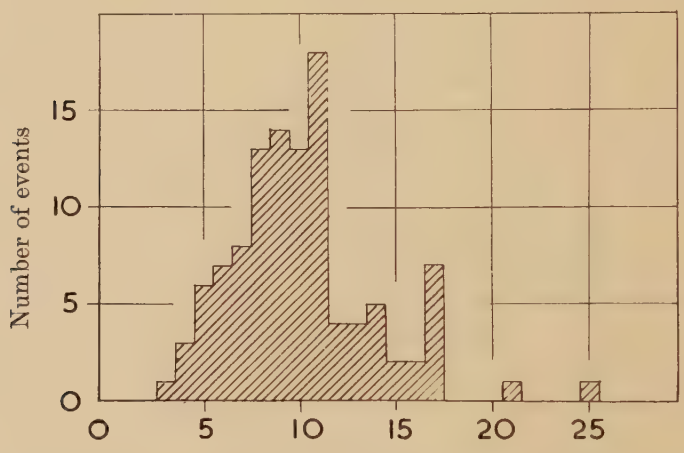
(*a*) Particles with $\bar{D}_{obs} \geqq 1{\cdot}60\ \mu$.

(*b*) Particles with $\bar{D}_{obs} \leqq 1{\cdot}35\ \mu$.

from the end of the tracks, is counted and the result represented in fig. 4 (*a*) and (*b*). The displacement of the maxima of the two histograms is immediately apparent. It is reasonable to assume that nearly all particles in group (*a*) are protons. This gives a mean value of $10{\cdot}8 \pm 0{\cdot}3$ delta rays in the last $1400\,\mu$ of a proton track. The corresponding value for a deuteron track may be determined in the following way :

The number of delta rays in the residual range R of a particle of unit charge and of mass M is M times the number of delta rays in the residual range R/M of a proton. By observing the average number of delta rays in the last $700\,\mu$ of proton tracks we obtain $6{\cdot}9 \pm 0{\cdot}3$ as the number of delta rays in the last $1400\,\mu$ of a deuteron track. These values were used for a statistical analysis of the histogram, fig. 5, representing delta ray counts on all the particles in fig. 3. The relative frequency of protons and deuterons thus obtained is included in table 1. The agreement with the result obtained by the magnetic deflection experiment is satisfactory. An internal consistency therefore exists between the three methods, which justifies the present scattering method.

Fig. 5



Histogram showing the distribution in the total number of delta rays, containing 4 grains or more, in $1400\,\mu$ range from the end of the tracks represented in fig 3.

§ 5. Discussion

A χ^2-test of the histogram, fig. 2, gives that the majority of the particles must be a single isotope of helium. The theoretical value for the mean scattering makes it reasonable to assume these to be alpha particles. The actual number of isotopes differing from ${}^4_2\mathrm{He}$ in the present material is difficult to estimate. ${}^6_2\mathrm{He}$ is beta radioactive with a period of 0·85 sec. However, no secondary particles, such as electrons, were observed to be emitted at the point where the helium particles came to rest in the emulsion. It is therefore reasonable to assume that the frequency of emission of energetic ${}^6_2\mathrm{He}$-nuclei is negligible.

The existence of energetic $^3_2$He isotopes among the disintegration products of a star is probable, however, for the following reasons. Firstly, tritons have been identified among the fast particles ejected from nuclear explosions (Camerini and Fowler, private communication), and their mechanism of ejection is probably very similar to that of the ejection of $^3_2$He nuclei. Secondly, the $^3_2$He/$^4_2$He ratio in meteorites ($\sim$0·3) is very much higher than that observed in minerals on the earth ($\sim 10^{-7}$), cf. Paneth, Reasbeck and Mayne (1952) and Singer (1952). This is commonly explained by the effect of cosmic radiation on the meteorites in space, in particular by the emission of low energy helium particles in nuclear evaporation.

A statistical analysis of the present sample of helium particles, assuming no $^6_2$He, shows that it contains less than 5% of $^3_2$He-nuclei. To obtain more exact figures of the relative frequency of the isotopes it is necessary to increase the number of observations, and to compare our result with that obtained from artificially accelerated alpha particles.

The large proportion of alpha particles may have a bearing on the problem of the emission of energetic helium nuclei. In order to avoid more speculative assumptions, for instance long range nuclear forces, etc. the following purely statistical interpretation of the mechanism of ejection has been suggested :

The average kinetic energy of the helium particles in the present investigation is of the order of magnitude $\sim$80 Mev. Thus, any mechanism of ejection based on nuclear evaporation processes may be excluded. However, the kinetic energy per nucleon is $\sim$20 Mev, i.e. of the order of magnitude of the zero point kinetic energy of nucleons inside a nucleus. Therefore, a group of nucleons may be ejected at large angles to one another during a nuclear explosion, and still be bound together forming a heavy fragment, provided the particles are ejected within a time interval of the order of magnitude of the ' characteristic nuclear time '. This may be defined as the nuclear radius divided by some average velocity of the nucleons, i.e. about $\sim 10^{-22}$ sec. However, this very simple mechanism would give a higher probability of emission of $^3_2$He compared with $^4_2$He, in contradiction to the results of the present experiment. It is therefore more reasonable to assume that the alpha particles are formed within the nucleus before they are knocked out by the primary particle producing the disintegration, or by one of the nucleons recoiling from the primary particle.

APPENDIX

Determination of the 'Noise' Level

The observed mean deviation $\bar{D}_{\text{obs}}$ of a track depends on the real deviation (theoretical) $\bar{D}$ due to multiple Coulomb scattering and the 'noise' $\bar{D}_{\text{noise}}$ due to spurious scattering. Assuming these two factors

to be statistically independent and following a normal distribution, we have

$$\bar{D}^2_{\text{obs}}=\bar{D}^2+\bar{D}^2_{\text{noise}}. \qquad (16)$$

In the first approximation $\bar{D}_{\text{noise}}$ is independent of cell size. Its magnitude was determined in the following way.

From the present material five He-tracks of range $>7000\,\mu$ were chosen at random. The mean deviation was measured for all the tracks separately, using the following five sets of varying cell sizes on each track :

$$t_1=t_0/4,\quad t_2=t_0/2,\quad t_3=3t_0/4,\quad t_4=t_0,\quad t_5=3t_0/2 \qquad (17)$$

where $t_0=t_0(R)$ is the varying cell size represented in fig. 1. The result of the measurements is given in table 2.

Table 2. $\bar{D}_{\text{obs}}$ in μ for Five Different Helium Tracks and Five Sets (17) of Varying Cell Sizes

	t_1	t_2	t_3	t_4	t_5
He (I)	0·210	0·370	0·590	0·85	1·70
He (II)	0·167	0·325	0·563	0·84	1·58
He (III)	0·191	0·375	0·650	0·97	1·69
He (IV)	0·175	0·370	0·540	0·86	1·42
He (V)	0·166	0·330	0·520	0·77	1·34
Mean	0·182	0·354	0·573	0·86	1·55

The theoretical mean deviation $\bar{D}_n$ corresponding to a particular t_n is related to the cell size t_m and the corresponding theoretical mean deviation $\bar{D}_m$ by the relation

$$\bar{D}_n=(t_n/t_m)^{3/2}\bar{D}_m. \qquad (18)$$

From pairs of equations of the type

$$\bar{D}_n{}^2{}_{\text{obs}}=\bar{D}_n{}^2+\bar{D}^2_{\text{noise}}, \qquad (19\,a)$$

$$\bar{D}_m{}^2{}_{\text{obs}}=\bar{D}_m{}^2+\bar{D}^2_{\text{noise}} \qquad (19\,b)$$

in connection with (18) and the table 2 we obtain values of the ' noise level ' $\bar{D}_{\text{noise}}$ differing slightly because of statistical fluctuations. The mean of the different values gives

$$\bar{D}_{\text{noise}}=(0{\cdot}19\pm0{\cdot}01)\mu.$$

Acknowledgments

The authors would like to express their gratitude to Professor J. Holtsmark and Professor R. Tangen for offering the laboratory facilities to carry out this work. Most of the plates used in the present investigation were supplied by H. H. Wills Physical Laboratory, University of Bristol, for which we wish to thank Professor C. F. Powell, F.R.S. We are indebted to the Royal Norwegian Council for Scientific and Industrial Research for financial support.

References

Crussard, 1952, *Thèse*, Paris.
Dainton and Fowler, 1953, in course of publication.
Daniel, Davies, Mulvey and Perkins, 1952, *Phil. Mag.*, **43**, 753.
Fowler, 1950, *Phil. Mag.*, **41**, 169.
Franzinetti, 1950, *Phil. Mag.*, **41**, 86.
Franzinetti and Payne, 1948, *Nature, Lond.*, **161**, 735.
Goldschmidt-Clermont, King, Muirhead and Ritson, 1948, *Proc. Phys. Soc.*, **61**, 183.
Holtebekk, 1951, *Hovedoppgave i fysikk*, University of Oslo.
Menon and Rochat, 1951, *Phil. Mag.*, **42**, 1232.
Molière, 1947, *Zeits. f. Naturforschung*, **2***a*, 133 ; 1948, *Ibid.*, **3***a*, 78.
Paneth, Reasbeck and Mayne, 1952, *Geochimica et Cosmo-chimica Acta*, **2**, 300.
Perkins, 1950, *Proc. Roy. Soc.* A, **203**, 399.
Singer, 1952, *Nature, Lond.*, **170**, 728.

CIX. CORRESPONDENCE.

Slip Lines and Etchpits

By S. Amelinckx
Lab. voor Kristalkunde, Rozier, 6, Gent, Belgium*

[Received July 6, 1953]

Recently experimental methods have been worked out to find correlations between distinct screw dislocations and slip lines. These methods are based upon the fact that the presence of a screw dislocation can be deduced from the occurrence of either a Frank spiral (Griffin 1952, Forty and Frank 1953) or a spiral pit (Amelinckx and Votava 1953).

To test any theory concerning the glide process in metals it would be interesting to know the distribution of dislocations along slip lines. Some information concerning this has already been obtained by Castaing and Guinier (1949) who considered the precipitation along slip lines in an Al-alloy (Cu 4%).

We have now developed another simple approach towards the study of the distribution of dislocations along slip lines.

It has been shown recently that the points of emergence of screw dislocations are preferentially attacked by chemical reagents, and that etchpits will be developed there (Gevers, Amelinckx and Dekeyser 1952, Gevers 1953, Horn 1952). If on the other hand the picture for glide, as developed by Frank and Read (1950) is correct, rows of screw dislocations should emerge along slip lines and by etching a row of etchpits should be developed.

To see whether this is confirmed by observation, we prepared specimens of high purity aluminium (99·99%). Relatively large crystals were developed by recrystallizing the specimens at about 600°c during 48 hours after having strained them about 2%. The specimens were then polished electrolytically using the De Sy–Haemers electrolyte (De Sy and Haemers 1941) (Perchloric acid (20%) : 1 vol. ; Aethyl Alcohol : 4 vol.). After having been deformed very gently, so that only widely spaced, apparently single slip lines were faintly visible, the specimens were very slightly etched with a mixture of aqua regia and hydrofluoric acid (50% hydrochloric acid ; 47% nitric acid ; 3% hydrofluoric acid). It was found that three main positions for etchpits (dislocations) can be distinguished :

(1) along slip lines (forming etch channels) (see photographs, Plate 42) ;

(2) along *certain* grain boundaries, depending upon the relative orientation of the two grains ;

(3) apparently at random.

---

* Communicated by W. Dekeyser.

No detailed analysis of the distribution of dislocations is attempted at this stage of the investigation, as the exact relation between the number of etchpits and the number of dislocations is not yet known.

From these observations we wish only to conclude, for the moment, that slip lines are indeed characterized by the occurrence of rows of dislocations along them, thus providing evidence for the Frank and Read mechanism. More striking is the fact that dislocations are not distributed uniformly along a slip line, but are clustered as can be seen on all photographs. This is what is to be expected on the basis of Mott's theory of work hardening (Mott 1952) for face centered metals.

It may be worth comparing these results on Al with observations made on ZnS (in the press). Although there is of course quite a difference in what concerns the binding and also the state of perfection of the two crystals, a certain similarity can be seen.

This work is part of a research scheme supported by I.R.S.I.A. (C.E.S.). I wish to thank Professor Dr. W. Dekeyser for his kind interest.

References

Amelinckx, S., and Votava, E., 1953, *Naturw.*, **40,** 290.
Castaing, R., and Guinier, A., 1949, *C.R. Acad. Sci., Paris*, **228**, 2033.
De Sy, A. L., and Haemers, H., 1941, *Stahl und Eisen*, **61,** 185.
Forty, A. J., and Frank, F. C., 1953, *Proc. Roy. Soc.* A, **217,** 262.
Frank, F. C., and Read, T. W., 1950, *Phys. Rev.*, **79,** 722.
Gevers, R., 1953, *J. Chim. Phys.* (in the press).
Gevers, R. Amelinckx, S., and Dekeyser, W., 1952, *Naturw.*, **39,** 448.
Griffin, L. J., 1952, *Phil. Mag.*, **43,** 827.
Horn, F. H., 1952, *Phil. Mag.*, **43,** 1210.
Mott, N. F., 1952, *Phil. Mag.*, **43,** 1151.

An Unusual Ω^0-Decay

By D. D. Millar and D. I. Page

The Physical Laboratories, The University, Manchester, 13*

[Received August 7, 1953]

In the course of a cloud chamber experiment on V-events carried out at Jungfraujoch an unusual example of a slow Ω^0-decay has been recently observed (Plate 43).†

The event is so far unique among cloud chamber observations in that the heavy positive secondary comes to the end of its range in the gas of the chamber. No decay product is emitted by the stopped particle, thus indicating that it is stable and so could be a proton.

---

* Communicated by P. M. S. Blackett, F.R.S.

† Following the discussion at the Bagnères Conference we use the symbol Ω^0 for the particle hitherto designated by V_1^0.

The heavily ionizing product has a range of $9{\cdot}1 \pm 0{\cdot}2$ cm in argon at 80 cm Hg and 18°C. The momentum of a proton with such a range is $0{\cdot}73 \times 10^8$ ev/c, with error $\sim 2\%$. The momentum of the negative particle, determined by the measured curvature in the magnetic field of 5 200 gauss, is $1{\cdot}26^{+0{\cdot}125}_{-0{\cdot}095} \times 10^8$ ev/c. The large error is caused by abnormal distortion conditions in the chamber at the time. Its ionization is estimated as less than twice the minimum value and it may be either a π- or a μ-meson. The total angle enclosed by the two decay tracks is $39° \pm 2°$. Assuming the decay scheme, $\Omega^0 \rightarrow P + \pi^- + Q$ (Armenteros *et al.* 1951) the calculated Q-value is then $34{\cdot}5^{+6{\cdot}5}_{-4{\cdot}5}$ Mev, in agreement with values obtained by other workers (Bagnères Conference 1953).

The momentum of the primary Ω^0 is $\sim 1{\cdot}9 \times 10^8$ ev/c and in the centre of mass system of the Ω^0 the proton is emitted at an angle of $\sim 160°$ to the line of flight of the primary. The lifetime of the particle in its rest system between the point of production in the lead above the chamber, as defined by two associated shower particles, and the point of decay is $\sim 2 \times 10^{-9}$ sec, of which 1×10^{-9} sec is spent within the chamber.

We are indebted to the administration of the Hochalpine Forschungsstation, Jungfraujoch, Switzerland and to Herr H. Wiederkehr for the facilities which make our work there possible, also to our colleagues of the Jungfraujoch team and to Mr. A. H. Chapman for assistance in running the apparatus.

References

Armenteros, R., Barker, K. H., Butler, C. C., and Cachon, A., 1951, *Phil. Mag.*, **42**, 1113.

Report of the Congrès International sur le Rayonnement Cosmique, 1953, to be published.

Fig. 1

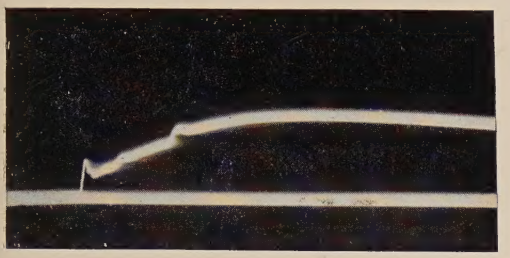
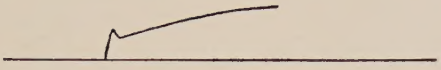


Fig. 1 (*a*)



r=16 km. i_0=12 kA.

Fig. 2

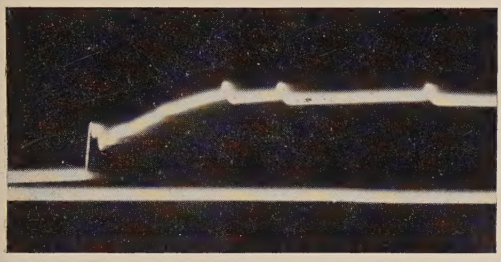


Fig. 2 (*a*)

r=20 km. i_0=23 kA.

30 v/m.

Fig. 3

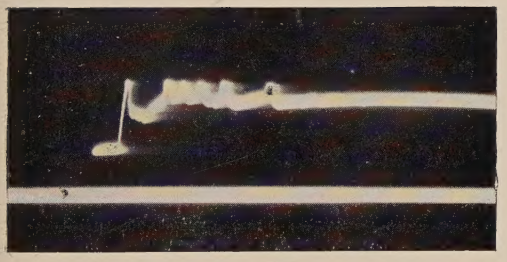
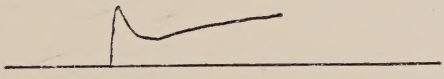


Fig. 3 (*a*)



r=28 km. i_0=51 kA.

Fig. 4

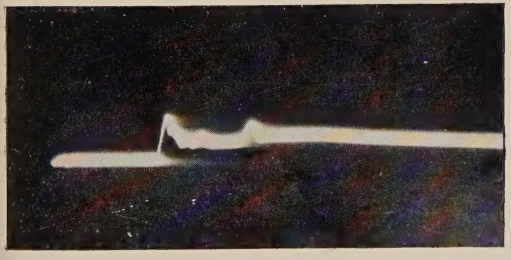
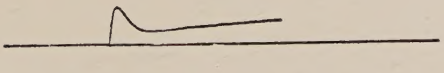


Fig. 4 (*a*)



300 μsec.

r=33 km. i_0=54 kA.

Fig. 5

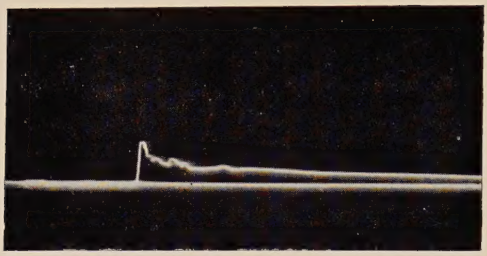
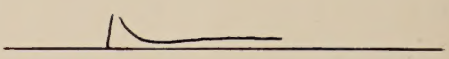


Fig. 5 (*a*)



r=47 km. i_0=16 kA.

Fig. 6

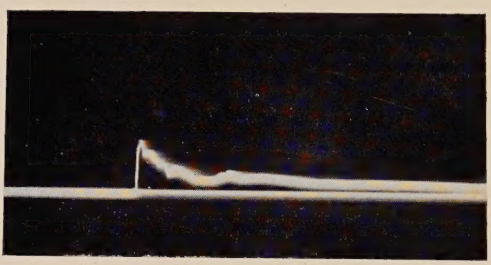
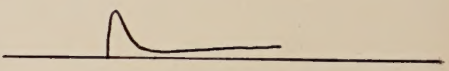


Fig. 6 (*a*)



r=53 km. i_0=24 kA.

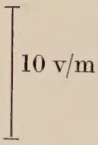


Fig. 7

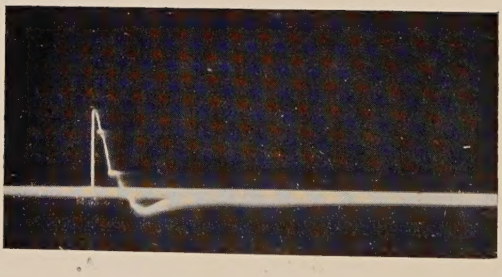


Fig. 7 (*a*)

r=100 km. i_0=80 kA.

Fig. 8

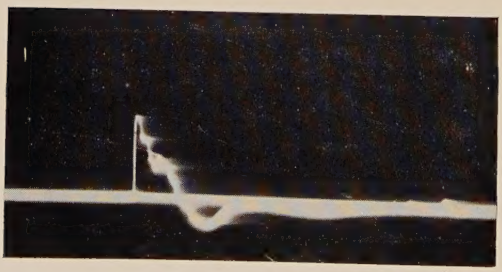
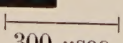
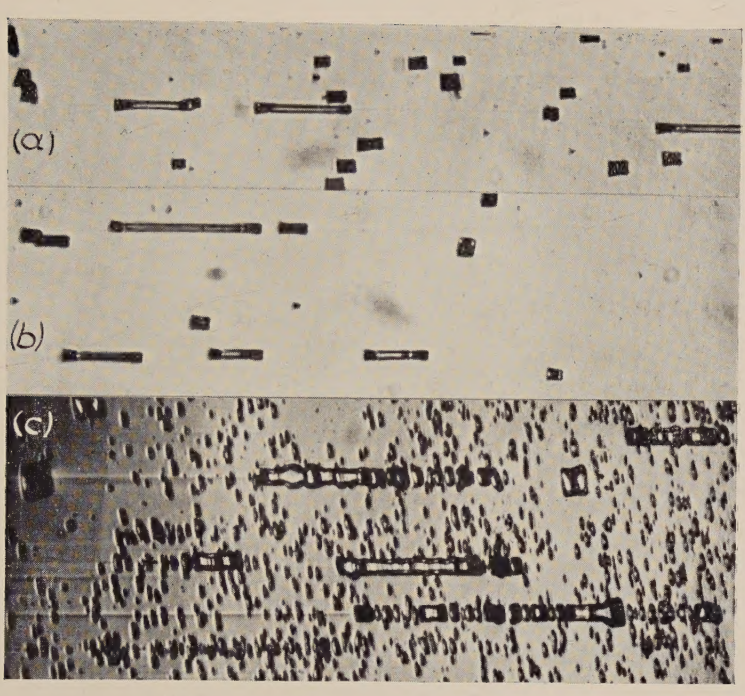
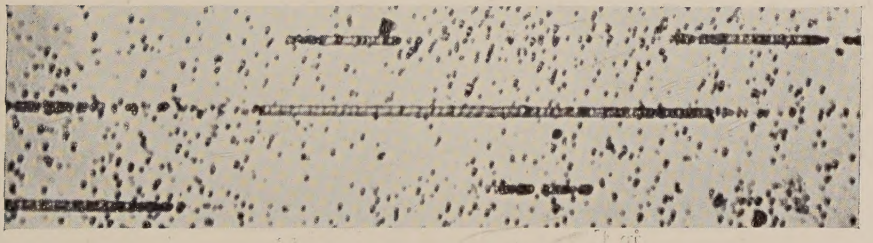


Fig. 8 (*a*)

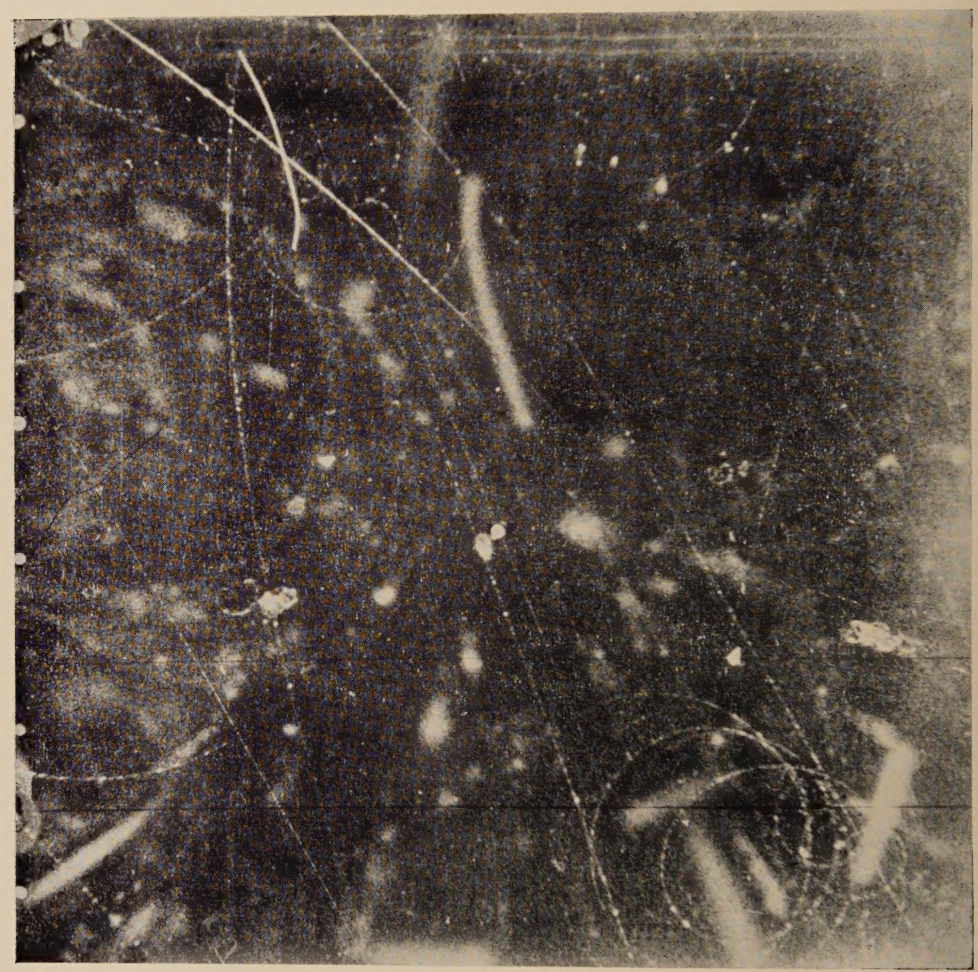


r=100 km. i_0=80 kA.





Etchpits developed along slip lines, on different faces of crystals of pure aluminium obtained by recrystallization. (×1200)



The decay of a Ω^0 particle, producing a stable positive particle which stops in the gas of the cloud chamber.